Methods for Addressing Uncertainty and Variability to Characterize Potential Health Risk from Trichloroethylene-Contaminated Ground Water at Beale Air Force Base in California: Integration of Uncertainty and Variability in Pharmacokinetics and Dose-Response

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

K. T. Bogen

May 24, 2001

Methods for Addressing Uncertainty and Variability to Characterize Potential Health Risk from Trichloroethylene-Contaminated Ground Water at Beale Air Force Base in California:

Integration of Uncertainty and Variability in Pharmacokinetics and Dose-Response

K.T. Bogen

Lawrence Livermore National Laboratory,
University of California, Livermore, CA*

May 24, 2001

*Address correspondence to: Dr. K.T. Bogen, Health and Ecological Assessment Div. L-396, Lawrence Livermore National Laboratory, Livermore, CA 94550-9234, USA, Tel: (925) 422-0902, Fax: (925) 424-3255, NET: bogen@LLNL.gov.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

v

ABBREVIATIONS and NOTATION

AM	arithmetic mean
$B_{MA,P}$	normalized MA-specific biologically effective (TCE or TCA) dose by pathway P
B(a,b)	Euler beta function of parameters a and b; equal to $\Gamma(a)\Gamma(b)/\Gamma(a+b) = \int_0^1 s^{a-1}(1-s)^{b-1}ds$.
B(x,a,b)	Incomplete beta function, equal to $\int_0^x s^{a-1}(1-s)^{b-1}ds$, for $0 \le x \le 1$.
cdf	cumulative probability distribution function
C	cytotoxic
C_x	concentration of x
CH	chloral hydrate
CV	coefficient of variation = SD/AM
CVM	CV of the AM
der	dermal
df	degrees of freedom
D_C	chronic effective TCA dose
D_{Ca}	acute effective TCA dose
$D_{MA,P}$	MA-specific biologically effective (TCE or TCA) dose by pathway P
DCA	dichloroacetic acid
DCVC	S-(1,2-dichlorovinyl)-L-cysteine
EX	a specified expectation with respect to the variate X (i.e., $\langle X \rangle$ or $\overline{X}$)
E_x	exposure pathway x
$f_x(y)$	probability density function pertaining to x, evaluated at y
$F_x(y)$	cdf pertaining to x, evaluated at y
G	genotoxic
GM	geometric mean
GSD	geometric standard deviation
ing	ingestion
inh	inhalation
J	a JUV (i.e., jointly uncertain and heterogeneous) variate
JUV	joint uncertainty and (interindividual) variability; of or pertaining to JUV
K_x	a constant pertaining to x
LN	lognormal
LN(μ,σ)	lognormally distributed with GM = e^μ and GSD = e^σ

LTWA	lifetime time-weighted average
MA	mechanistic assumption, i.e., assumed mode/mechanism of action
$N(\mu,\sigma)$	normally distributed with AM $= \mu$ and SD $= \sigma$
NRC	National Research Council
NTP	National Toxicology Program
P	pathway or route of TCE exposure (intake)
P or p	probability
p_{adj}	adjusted probability
PBPK	physiologically based pharmacokinetic
r	product-moment correlation
R	increased individual lifetime probability of incurring a toxic (cancer and/or noncancer) endpoint due to TCE exposure
SD	standard deviation
SDM	SD of the AM
t	time, or reference to Student's t-distribution
TCA	trichloroacetic acid
TCE	trichloroethylene
USEPA	U.S. Environmental Protection Agency
U_x	an uncertain variate pertaining to x
V_x	a "heterogeneous" variate pertaining to x; i.e., different variate values pertain to different individuals at risk
VOC	volatile organic compound
X	vector of uncertain and/or heterogeneous variates, $\{X_1, X_2, ... \}$
$\langle X \rangle$	expectation with respect to uncertainty pertaining to the variate X
$\overline{X}$	expectation with respect to interindividual variability pertaining to the variate X
$X_{MA,P}$	MA-specific TCE exposure (intake) by pathway P
$\beta(x,a,b)$	Incomplete beta function ratio, or regularized incomplete beta function, equal to $B(x,a,b)/B(a,b)$, for $0 \le x \le 1$.
$\Phi(z)$	cdf of a standard normal (Gaussian) variate Z, $= \mathrm{Prob}(Z \le z)$
$\rho_r(x,y)$	rank correlation coefficient between variates x and y

ABSTRACT

Traditional estimates of health risk are typically inflated, particularly if cancer is the dominant endpoint and there is fundamental uncertainty as to mechanism(s) of action. Risk is more realistically characterized if it accounts for joint uncertainty and interindividual variability within a systematic probabilistic framework to integrate the joint effects on risk of distributed parameters of all (linear as well as nonlinear) risk-extrapolation models involved. Such a framework was used to characterize risks to potential future residents posed by trichloroethylene (TCE) in ground water at an inactive landfill site on Beale Air Force Base in California. Variability and uncertainty were addressed in exposure-route-specific estimates of applied dose, in pharmacokinetically based estimates of route-specific metabolized fractions of absorbed TCE, and in corresponding biologically effective doses estimated under a genotoxic/linear (MA_G) vs. a cytotoxic/nonlinear (MA_C) mechanistic assumption for TCE-induced cancer. Increased risk conditional on effective dose was estimated under MA_G based on seven rodent-bioassay data sets, and under MA_C based on mouse hepatotoxicity data. Mean and upper-bound estimates of combined risk calculated by the unified approach were $<10^{-6}$ and $<10^{-4}$, respectively, while corresponding estimates based on traditional deterministic methods were $>10^{-5}$ and $>10^{-4}$, respectively. It was estimated that no TCE-related harm is likely to occur due to any plausible residential exposure scenario involving the site. The systematic probabilistic framework illustrated is particularly suited to characterizing risks that involve uncertain and/or diverse mechanisms of action.

1. INTRODUCTION

This report describes methods and results pertaining to Phase 2 of a study involving quantitative consideration of joint uncertainty and interindividual variability in risk to hypothetical future residents posed by trichloroethylene (TCE) in ground water at the inactive landfill Site LF-13 on Beale Air Force Base in California. The background of this study is discussed below, followed by summaries of: the rationale for this study's focus on quantitative analysis of joint uncertainty and variability; technical hurdles posed by undertaking such an analysis in a way that explicitly addresses carcinogenic dose-response of TCE in view of fundamental uncertainty concerning its carcinogenic mode of action; and study goals of Phase 2 of the analysis undertaken of risk posed by TCE at Site LF-13. Specific methods used to address the latter goals are presented in Section 2 of this report. Results obtained by applying these methods are presented in Section 3, followed in Section 4 by a discussion of the results obtained. References cited in this report are listed in Section 5. Appendix 1 supplies mathematical details concerning the "method of moments" used throughout in this report to make assumptions about lognormal variates. Appendix 2 supplies a proof concerning the distribution of a linear function of correlated t-distributed variates. Finally, Appendix 3 documents all calculations performed for this study.

The general background of the present study and its Phase-1 counterpart is provided in Section 1.1 below, followed by: a summary of the rationale for the emphasis in this report placed on quantitative analysis of joint uncertainty and variability (Section 1.2), a discussion of the present fundamental uncertainty pertaining to mechanism(s) of action for TCE-induced cancer (Section 1.3), issues involving quantitative analysis of joint uncertainty and variability in dose-response for TCE-induced cancer (Section 1.4), and the specific goals of the present report (Section 1.5).

1.1. Background

Traditional point estimates of risk are calculated deterministically using worst-case assumptions for some or all input parameters, in a way that does not quantitatively account for uncertainty and interindividual variability pertaining to these parameters. Traditional point-estimates of risk are thus typically inflated and health-conservative, particularly if the cancer is the dominant endpoint and there is fundamental uncertainty

4

as to mechanism(s) of action. Risk is more realistically characterized if it accounts for joint uncertainty and interindividual variability within a systematic probabilistic framework to integrate the joint effects on risk of distributed parameters in all (linear as well as nonlinear) risk-extrapolation models for all (cancer as well as noncancer) endpoints involved. Because no such systematic probabilistic framework existed, one was developed for the present case study involving the inactive Landfill Site LF-13 on Beale Air Force Base in California, where groundwater contaminated with trichloroethylene (TCE) has moved beyond the site boundary. Soil-vapor extraction and air-stripping treatment of groundwater have been undertaken to reduce concentrations of TCE and other volatile organic compounds in ground water beneath Site LF-13 (URSGWC, 1998). Site LF-13 is located in a currently rural area of the Sacramento Valley of California, where groundwater wells are the principle source of domestic water supplies. The present analysis was undertaken to provide a realistic characterization of hypothetical TCE-related risks associated with potential future domestic/residential uses of groundwater from beneath Site LF-13, in view of the possibility that residential populations may eventually occupy lands adjacent to the site.

This study was conducted in two phases. Phase 1 focused on the impact of joint uncertainty and interindividual variability (JUV) on estimates of combined TCE intakes via different exposure pathways (Daniels et al., 1999, 2000). Uncertainty here refers to an absence of measurement data or incomplete knowledge; interindividual variability (or "variability") here refers to true differences or heterogeneity in an empirical, risk-related characteristic (e.g., physiological differences) among individuals in a population (Bogen and Spear, 1987; NRC, 1994). Although results of the Phase 1 analysis were presented as a characterization of risk rather than exposure, risk was estimated in that analysis simply as the product of estimated combined exposures (in mg kg^{-1} d^{-1}) and respective factors representing carcinogenic potency (in kg d mg^{-1}), each of which factor was represented by a single point estimate. Thus, JUV in risk characterized in Phase 1 reflected only JUV in estimated exposure, and in no way addressed JUV associated with TCE pharmacokinetics, dose-response, alternative mechanisms of toxic action, or multiple toxic endpoints. The TCE concentration in Phase 1 was estimated based on groundwater-monitoring data for a well on Site LF-13 near the possible location of a future groundwater extraction and distribution system (Purrier, 1997). After

considering concentration uncertainty and JUV in potential multi-route exposures to TCE from Site LF-13 ground water, corresponding JUV in risk was characterized and compared to corresponding risk estimators that were calculated using traditional deterministic methods (Daniels et al., 2000).

Phase 2 of the study described above is the subject of the present report. Phase 2 involved the development of new methods allowing additional information to be integrated into a Phase-1-type TCE risk assessment for Site LF-13. This additional information involves JUV in predicted risk *conditional on* route-specific TCE exposures. As further explained below, this was accomplished by combining exposure distributions and methods presented in the Phase-1 study with TCE-related pharmacokinetic and dose-response methods and information developed in the present study, to provide an improved characterization of TCE-related risk associated with Site LF-13 at Beale AFB.

1.2. Importance of Quantitative Analysis of Joint Uncertainty and Variability

This study focuses on integrating information on joint uncertainty and interindividual variability (JUV) to obtain more meaningful and more realistic estimates of exposure and risk. In the report, *Science and Judgment in Risk Assessment*, the National Research Council (NRC) emphasized the importance of distinguishing clearly between uncertainty (i.e., lack of knowledge) and interindividual "variability" (i.e., heterogeneity or differences pertaining to people at risk) in risk assessment (NRC, 1994). Uncertainty in characterized risk reflects the extent to which a risk estimate is likely to be erroneous, due to gaps in data and/or theory that imply statistical and/or model-specification error. Interindividual variability in characterized risk reflects the extent to which a risk is unequally imposed on members of the population at risk. While uncertainty reduces the confidence or reliability that can be placed in a risk estimate, variability can be viewed as a measure of perceived unfairness or inequity represented by the distribution of imposed risks. Because reliability and equity issues are clearly related to perceived and/or statutorily defined risk acceptability criteria, both these dimensions may be relevant to risk-management policy decisions.

Quantitative characterization of joint uncertainty and variability (JUV) in risk is a way to address risk-related uncertainty and variability concisely and explicitly to

facilitate risk management decisions. When JUV is addressed quantitatively in the input distributions used to characterize the inputs (e.g., on ambient concentration, uptake, and dose-response) of a risk assessment, the distinction between uncertainty and variability ought to be maintained rigorously throughout the analytic process so that uncertainty and variability can be reflected distinctly in the calculated risk. This recommendation was expressed by the NRC (1994, p. 242) as follows:

> "A distinction between uncertainty (i.e., degree of potential error) and inter-individual variability (i.e., population heterogeneity) is generally required if the resulting quantitative risk characterization is to be optimally useful for regulatory purposes, particularly insofar as risk characterizations are treated quantitatively. The distinction between uncertainty and individual variability ought to be maintained rigorously at the level of separate risk-assessment components (e.g., ambient concentration, uptake, and potency) as well as at the level of an integrated risk characterization."

If no distinction is made between uncertainty-related and heterogeneity-related distributions associated with inputs to a given risk calculation, then the resulting distribution necessarily reflects risk to an individual selected at random from the exposed population (Bogen and Spear, 1987). By definition, this resulting distribution cannot be used for any regulatory decision intending to address equity issues by focusing on risk borne by relatively more sensitive and/or relatively more highly exposed members of the population at risk. Another advantage of distinguishing between uncertainty and variability is that it permits one to estimate the uncertainty in the risk to the individual who is "average" with respect to all characteristics that are heterogeneous among individuals at risk. Only the latter quantity can be used to estimate corresponding uncertainty in predicted population risk (i.e., uncertainty in the predicted number of cases), and thus, in particular, to estimate the likelihood of zero cases (i.e., the likelihood that remediation of the exposure scenario considered will have no positive impact whatsoever on public health) (Bogen and Spear, 1987).

1.3. Uncertainty in Mechanism(s) of Toxic Action

Liver is clearly a target tissue for TCE-induced cancer based on lifetime bioassay data on chronically exposed mice; relatively large acute, subchronic, or chronic TCE exposures are hepatotoxic in multiple species; and hepatocellular toxicity in mice is the most sensitive TCE-induced noncancer (but possibly cancer-related) endpoint (Bogen

7

and Gold, 1997; Bogen et al., 1988; USEPA, 1985). Limited epidemiological data also support liver cancer as a TCE-induced endpoint in occupationally exposed humans (Wartenberg, 2000). DNA-binding and weak mutagenicity associated with TCE metabolites after TCE administration indicates that genotoxicity may be responsible for some or all TCE-induced cancer (Bogen and Gold, 1997; Fahrig et al., 1995). Two TCE metabolites in particular, trichloroacetic acid (TCA) and dichloroacetic acid (DCA), both induce and promote liver tumors in a mouse strain (B6C3F1) which is positive for TCE-induced liver cancer, whereas liver tumors did not appear in rats exposed to either TCE by gavage or to TCA via drinking water (Bogen and Gold, 1997; Bull et al., 1990; DeAngelo et al., 1997; DeAngelo et al., 1991; Herren-Freund et al., 1987; Pereira, 1996; Pereira and Phelps, 1996). DCA in particular was found recently to be weakly mutagenic in mouse lymphoma cells with a mutagenic potency similar to the classic mutagen ethyl methanesulfonate, whereas only very weak mutagenic activity was detected using either the major reactive TCE metabolite, chloral hydrate (CH), or its breakdown product TCA (Harrington-Brock et al., 1998; Moore and Harrington-Brock, 2000). Initial studies found DCA to be more reactive and toxic than TCE, and thus more likely to account for observed TCE-induced cancer in bioassay mice (Larson and Bull, 1992a-b; Templin et al., 1993). However, more recent studies that controlled for *ex vivo* formation of DCA during sample preparation indicate that very little, if any, DCA was actually produced in TCE-exposed B6C3F1 mice, imply the same for humans as well (Lash et al., 2000a), and are consistent with the hypothesis that DCA is unlikely to explain TCE-induced mouse tumors (Andersen et al., 1998; Merdink et al., 1998) or to be relevant in extrapolating TCE-induced cancer risk for humans (Bull, 2000).

The generally weak mutagenicity of TCE and its metabolites (Moore and Harrington-Brock, 2000), as well as correlations between hepatotoxic indicators induced by reactive TCE metabolites and precursors to TCE-induced liver tumorigenesis (Bogen and Gold, 1997), provide substantial (although not definitive) support a cytotoxic mechanism of TCE-induced carcinogenic action. Hepatotoxic lipid peroxidation was found to be induced by TCA in mice and rats, but mice were found to be more sensitive than rats (Larson and Bull, 1992a). This differential sensitivity to a TCA-induced cytotoxic endpoint is consistent with a cytotoxicity-based explanation of TCE-induced liver tumors in mice but not rats. A more recent study of lipid peroxidation induced in

B6C3F1 mouse liver concluded that the amount of such peroxidation induced by "TCA equaled that induced by CH, whereas that from [trichloroethanol, another major, but less toxic and reactive, TCE metabolite] was 3- to 4-fold lower, suggesting that metabolism of CH to TCA may be the predominant pathway leading to lipid peroxidation" (Ni et al., 1996). Lipoperoxidation-induced oxidative stress may explain or correlate with the induction of hepatocellular replicative DNA synthesis and hepatocellular proliferation that has been observed in TCA-exposed B6C3F1 mice (Dees and Travis, 1994). Increased cell proliferation, in turn, either alone or in combination with genotoxic conditions, has long been considered sufficient to explain increased rates of cancer in view of biologically based mechanistic cell-kinetic multistage cancer theory, as well as based on experimental, epidemiological and clinical observations (Ames and Gold, 1990a; Ames and Gold, 1990b; Ames et al., 1993,1995; Armitage and Doll, 1957; Bogen, 1989; Cohen and Ellwein, 1990,1991; Moolgavkar, 1983; Moolgavkar and Knudson, 1981; Moolgavkar et al., 1988).

Statistical considerations support rejecting lung as a significant target site for TCE-induced cancer in rodents (Bogen and Gold, 1997). Epidemiological evidence of TCE-induced kidney cancer in humans has been characterized variously as "limited", "suggestive", and "neither consistent nor convincing," but additional recent data appear to indicate that kidney cancer risk is elevated among metabolically susceptible workers with very large TCE exposures (Lash et al., 2000b). While National Toxicology Program (NTP) bioassay data also indicate that kidney cancer is induced experimentally by chronic TCE exposure is the rat (but not mouse), these data were all judged to be "inadequate" after NTP review, with mild to severe renal toxicity observed at every non-control dose level in every species/sex combination in the bioassays (Bogen and Gold, 1997; NTP, 1988,1990). The NTP rats studies and limited/suggestive epidemiological data comprise the only evidence that TCE-induced renal tumors are plausibly relevant to humans (Bogen and Gold, 1997; Bogen et al., 1988; USEPA, 1985; Lash et al., 2000b). The rat tumor data are consistent with a cytotoxic mechanism of action for renal carcinogenesis, although mutagenicity of renal TCE-metabolites such as S-(1,2-dichlorovinyl)-L-cysteine (DCVC) indicates that genotoxicity may also play a role (Bogen and Gold, 1997; Fahrig et al., 1995; Lash et al., 2000a-b). Interestingly, while subchronically administered TCA and acutely administered DCVC were both found to

be nephrotoxic and to induce cell proliferation in rat kidney tubules, the DCVD-induced response in mice (in which TCE-induced kidney tumors have not been observed) was much more pronounced than in rats (in which TCE-induced kidney tumors have been observed) (Acharya et al., 1997; Eyre et al., 1995). Consequently, the same uncertainties regarding the mechanism of TCE-induced hepatocarcinogenicity apply to the mechanism of possible TCE-induced renal tumors.

The U.S. Environmental Protection Agency (USEPA) has not explicitly endorsed the quantitative combination of "model" uncertainty with other types of uncertainty in cancer risk assessments for compounds like TCE (USEPA, 1996). This USEPA position is consistent with a recent NRC recommendation against this type of quantitative treatment as opposed to narrative/qualitative comparisons of model-specific analyses (NRC, 1994). However, the restriction of "model" uncertainty to be characterized only nonquantitatively needlessly reduces the clarity of risk analysis, because there is no logical merit to the distinction between "model" and "parameter" uncertainty. As the NRC report itself pointed out, the former is logically equivalent to the latter when incorporated into a suitably general model that specifies, through values assigned to one or more uncertain parameters, any particular but uncertain model characteristics (i.e., substructures) of concern (NRC, 1994; p. 187).

1.4. Technical Issues Posed by Quantitative Analysis of Joint Uncertainty and Variability in Dose-Response for TCE-Induced Risk

In view of the issues discussed above, there were several technical issues that had to be addressed in this study due to its primary focus on quantitative analysis of JUV in dose-response for TCE-induced risk. These issues concern the lack of coordinated methods that consistently and simultaneously address:

(i) multiple toxic (in this case, cancer and noncancer) endpoints with potentially disparate dose-response relations,

(ii) multiple plausible mechanisms of carcinogenic action,

(iii) efficient treatment of pharmacokinetic relations, and

(iv) integrated, quantitative treatment of JUV in exposure, dose-response, and risk calculations.

General approaches to issues (i) and (iv) have been reviewed (Bogen, 1995; NRC, 1994). Also pertaining to issues (i) and (iv) are proposed methods to extend quantitative

probabilistic methods now commonly applied in cancer risk assessment to noncancer endpoints, which involve replacing traditional uncertainty/safety factors by corresponding empirically based, or reasonable default, probability distributions (Baird et al., 1996; Carlson-Lynch et al., 1999; Dourson et al., 1996; Lewis, 1993; Renwick, 1993; Slob and Pieters, 1998; Weil, 1972). Issue (ii) is a major focus of the proposed USEPA guidelines for carcinogen risk assessment (USEPA, 1996), but in this regard the USEPA recommends a non-quantitative, narrative approach that cannot possibly address issue (iv). Concerning issue (iii), a number of physiologically based pharmacokinetic (PBPK) models have been developed for TCE (Abbas and Fisher, 1997; Allen and Fisher, 1993; Fisher and Allen, 1993; Fisher et al., 1998; Stenner et al., 1998), and corresponding methods for efficient PBPK analysis have been developed under different mechanistic assumptions concerning TCE-induced cancer (Bogen, 1988; Bogen and Gold, 1997). Concerning issues (ii) and (iii), PBPK methods for TCE have been applied under alternative mechanistic assumptions (Fisher and Allen, 1993), but this has never been done in a way that integrates JUV information or efficient analytic (as opposed to numerical) PBPK-calculation methods.

In recently proposed revised methods for deriving Ambient Water Quality Criteria, the USEPA indicated that a goal of these methods should be to integrate cancer and noncancer assessments, and more specifically "to harmonize cancer and noncancer dose-response approaches and permit comparisons of cancer and noncancer risk estimates" (USEPA, 1998; p. 59,97). To the extent this goal was achieved, these proposed methods would provide guidance on how to address issues (i)-(iv) in a systematic probabilistic framework for risk assessment. While the proposed methods do address multiple (cancer and noncancer) endpoints and alternative (linear vs. nonlinear) mechanisms of carcinogenic action, they do not specifically facilitate or even address their stated goal of integrating cancer and noncancer dose-response methods to yield comparable or aggregate measures of risk. Furthermore, this goal is unnecessarily impeded by some of the proposed methods, including those that either: address dose-response differently for noncarcinogens vs. "nonlinear" carcinogens, address generic pharmacokinetic considerations differently for noncarcinogens vs. ("linear" or "nonlinear") carcinogens, consider non-ingestive exposure as well as human interindividual variability in dose-response for noncarcinogens and "nonlinear"

11

carcinogens but not for "linear" carcinogens, or yield estimates of risk for "linear" carcinogens but do not for noncarcinogens and "nonlinear" carcinogens.

The impact of such inconsistencies on the problem of how to do unified risk assessment for cancer and noncancer endpoints is illustrated by the issue of whether or how to apply a toxicodynamic scaling factor to account for systematic interspecies differences in response as a function of biologically effective dose. For noncarcinogens, recently proposed USEPA methods include a good explanation of why, in the absence of relevant data, it is appropriate to apply two separate scaling factors (by default, each equal to a factor of 3) to account for interspecies toxicokinetic and toxicodynamic differences, respectively (USEPA, 1998; p. 140):

> "The rationale for the use of PBPK models is that the pharmacokinetics and pharmacodynamics of a chemical each contribute to a chemical's observed toxicity, and specifically, to observed differences among species in sensitivity. Pharmacokinetics describes the absorption, distribution, metabolism, and elimination of chemicals in the body, while pharmacodynamics describes the toxic interaction of the agent with the target cell. In the absence of specific data on their relative contributions to the toxic effects observed in species, each is considered to account for approximately one half of the variability in observed effects, as is assumed in the development of RfCs and RfDs [i.e., reference concentrations and doses, respectively]. The implication of this assumption is that an interspecies uncertainty factor of 3 rather than 10 could be used for deriving an RfD when valid pharmacokinetic data and models can be applied ... ."

For carcinogens, there is agreement that animal-to-human extrapolation of toxicokinetically equivalent effective dose may be accomplished by the use of an appropriate, validated PBPK model if one is available, and if not, by assuming that toxicokinetically equivalent doses scale proportional to body surface area or (body weight)$^{0.75}$ (USEPA, 1992,1996,1998). However, federal policy concerning how/whether to apply an interspecies toxico*dynamic* scaling factor is not consistent. For example, the Health/Risk Assessment Committee of the Integrated Chlorinated Solvents Project (a committee comprised of representatives from four federal agencies) held that "it is strongly arguable that the surface-area correction is not a correction on dose to allow for pharmacokinetics, but rather a correction on risk to allow for many factors, including pharmacodynamics" (USEPA, 1987a; p. 125). For "linear" carcinogens, however, the USEPA has more recently proposed that no interspecies

toxicodynamic scaling factor is required for carcinogens whenever a PBPK approach has been used to account for interspecies toxicokinetic differences (USEPA, 1998).

Likewise, interindividual variability in sensitivity/susceptibility *per se* to environmentally induced cancer is not typically considered in risk extrapolations for carcinogens assumed to have a genotoxic/linear-no-threshold mechanism of action (USEPA, 1996, 1998). In this respect, past practice has been to focus (implicitly) on risk to persons who have an average level of susceptibility, when there is no reason to predict that the exposed population is one that may reflect an unusual degree of hypersusceptibility to environmentally induced cancer (NRC, 1994). For noncarcinogens, however, a so-called "uncertainty" factor of up to 10 has traditionally been applied "to account for the variation in sensitivity (intraspecies variation) among the members of the human population"; and a similar factor was proposed recently by the USEPA for use with all "nonlinear" carcinogens (USEPA, 1998; p. 110,122). Confusion between uncertainty and interindividual variation also appears in proposed new approaches to model differences in human sensitivity by probabilistic methods rather than by the traditional use of "uncertainty" factors (Carlson-Lynch et al., 1999; Slob and Pieters, 1998). Because there is little doubt that substantial human variability exists in susceptibility to environmentally induced cancer (NRC, 1994), a truly systematic probabilistic framework to assessing risks pertaining to cancer and noncancer endpoints clearly requires a consistent approach to intraspecies variability in dose-response.

Recent proposals for so-called "unified" or "comprehensive" approaches to risk assessment for cancer and noncancer endpoints (Butterworth and Bogdanffy, 1999; Gaylor et al., 1999) fail to address the complete set of issues (i)-(iv) listed above. These proposals essentially recommend merely that a traditional safety-factor approach be used for cancer and noncancer endpoints alike; they focus on how to define exposure levels that protect against a single endpoint, rather than on how to calculate actual levels of aggregate risk for both cancer and noncancer endpoints. Therefore, no methods or studies exist that address the complete set of issues (i)-(iv) for integrated risk characterization.

1.5. Study Objectives

This study (Phase 2 together with Phase 1) was designed to accomplish two overall objectives. The first overall objective was to provide to the U.S. Air Force and regulatory agencies new quantitative procedures that address JUV in exposure and dose-response assessment to better characterize potential health risk. Such methods could be used at sites where populations may now or in the future be faced with using groundwater contaminated with low concentrations of TCE. The second overall objective was to illustrate and explain the application of these procedures with respect to available data for TCE in ground water beneath an inactive landfill site that is undergoing remediation at Beale Air Force Base in California. The results of this case study are intended to illustrate how the more realistic and more meaningful risk estimates obtained using methods we describe compare to corresponding conservative risk estimates calculated using a traditional deterministic screening-level approach. Application of the methods developed in this project can lead to more reasonable and equitable risk-acceptability criteria for potentially exposed populations at specific sites.

The specific objective of the present report is to describe consistent and coherent methods devised to address issues (a)-(d) discussed in the previous subsection, and to report and discuss an application of these methods, together with other methods and information developed in Phase 1 of this project, to the specific problem of characterizing risk posed by TCE in ground water at Site LF-13 at Beale Air Force Base.

2. METHODS

Methods used to address joint uncertainty and interindividual variability (JUV) in risk posed by TCE contamination at Site LF-13 was calculated and characterized as described below in subsections pertaining to: (1) the systematic probabilistic framework adopted for this analysis, (2) TCE concentration and route-specific exposures, (3) corresponding biologically effective doses and related physiologically based pharmacokinetic (PBPK) considerations, (4) dose-response for cancer and noncancer endpoints, (5) characterization of joint uncertainty and interindividual variability (JUV) in risk as a function of JUV in input parameters relating to topics (2)-(4) listed above, and (6) data analysis and computation.

Consistent with established JUV notation, an overbar (i.e., $^{-}$) here denotes expectation with respect to heterogeneous parameters only, angle brackets (i.e., $\langle\ \rangle$) denote expectation with respect to uncertain parameters only, each subscripted U denotes a corresponding purely uncertain variate, and each subscripted V denotes a corresponding purely interindividually heterogeneous variate (Bogen and Spear, 1987; NRC, 1994; Bogen, 1995). The operator E (e.g., as in EX) is used to signify expectation with respect either to variability or to uncertainty, bot not to both (e.g., to either the variability-expectation $\overline{X}$, or to the uncertainty-expectation $\langle X \rangle$, with respect to a variate X). Other italic capital letters are used to denote (potentially) JUV variates. Also, each subscripted K denotes a corresponding constant used below to estimate risk, $X_{MA,P}$ (in mg kg^{-1} d^{-1}) denotes a mechanistically relevant measure of TCE intake by the indicated exposure pathway (P) that pertains to the indicated mechanism/mode of action (MA) for TCE-induced toxicity, and $D_{MA,P}$ denotes a corresponding biologically effective dose, where, for both X and D, the subscript MA specifies either a genotoxic (G) or cytotoxic (C) assumed mode of action, and the subscript P indicates either an ingestion (ing), inhalation (inh), or dermal (der) exposure pathway. Additional related variates are defined below.

Some variates defined in Daniels et al. (2000), denoted as they were in that report, are also referred to below. All constants and variates defined in the present report used as inputs to estimate risk are defined below and are summarized in Table 2, which appears at the end of Methods (after Section 2.6) prior to Results (Section 3).

2.1. Systematic Probabilistic Framework

Health risk associated with residential exposure to TCE from ground water at Site LF-13 on Beale Air Force Base in California was analyzed using the systematic probabilistic framework summarized in Figure 1. Total associated risk, R, was defined as the increased individual lifetime probability of incurring a toxic (cancer and/or noncancer) endpoint due to TCE exposure from three pathways: direct ingestion of TCE-contaminated groundwater, dermal absorption of TCE while showering or bathing, and inhalation of TCE volatilized from water to household air. For volatile organic compounds (VOCs) such as TCE, these three pathways typically are the most significant contributors to total daily residential intake. Each route-specific TCE intake

was converted to a corresponding biologically effective dose for each MA and toxic endpoint considered, where this conversion was made using efficient MA-specific forms of a human PBPK model. Route-specific effective doses were summed for each MA to obtain (two) measures of MA-specific total effective dose (see Section 2.2). As detailed in Section 2.4, two MAs were considered for TCE: a genotoxic MA (MA_G) and a cytotoxic MA (MA_C), with both MAs considered potentially relevant to cancer risk posed by TCE exposure, but only MA_C considered relevant to noncancer risk posed by TCE exposure. Briefly, it was assumed that liver cytotoxicity is the most sensitive noncancer endpoint for TCE in humans based on the most sensitive experimental (mouse) data, that hepatotoxicity may (itself, or as the most sensitive available cytotoxic indicator) also explain and/or contribute to TCE-induced cancer observed in animal bioassays, and that genotoxicity may additionally explain and/or contribute to TCE-induced cancer observed in rodent bioassays. Increased likelihoods of cancer and of hepatotoxicity were each modeled as a MA-specific function of PBPK-based biologically effective dose in animals. Interspecies extrapolation of pharmacokinetic differences was obviated by consistent application of relevant PBPK models. Interspecies pharmacodynamic differences in dose-response were extrapolated using a single method applied to both cancer and noncancer endpoints. Intraspecies (interindividual) variability in human dose-response was modeled identically for both cancer and noncancer endpoints. Finally, increased risks of incurring either or both endpoints were estimated with respect to associated JUV, and these estimates were compared to corresponding traditional-type risk estimates obtained using deterministic methods.

The following subsections describe the specific methods used to apply the general approach just summarized, with respect to route-specific TCE exposures (Section 2.2), corresponding biologically effective doses (Section 2.3), dose-response for TCE-induced toxicity (Section 2.4), and unified risk characterization (Section 2.5).

2.2. TCE Exposure

Predicted rates ($X_{G,ing}$, $X_{G,inh}$, and $X_{G,der}$; mg kg^{-1} d^{-1}) of route-specific, lifetime time-weighted average (LTWA) exposure to TCE due to Site LF-13 groundwater contamination at Beale AFB were calculated from corresponding rates (E_{Ing}, E_{Inh}, and E_{Derm}) and associated JUV defined in Equations 1-3 of the Phase-1 report (Daniels et al.,

16
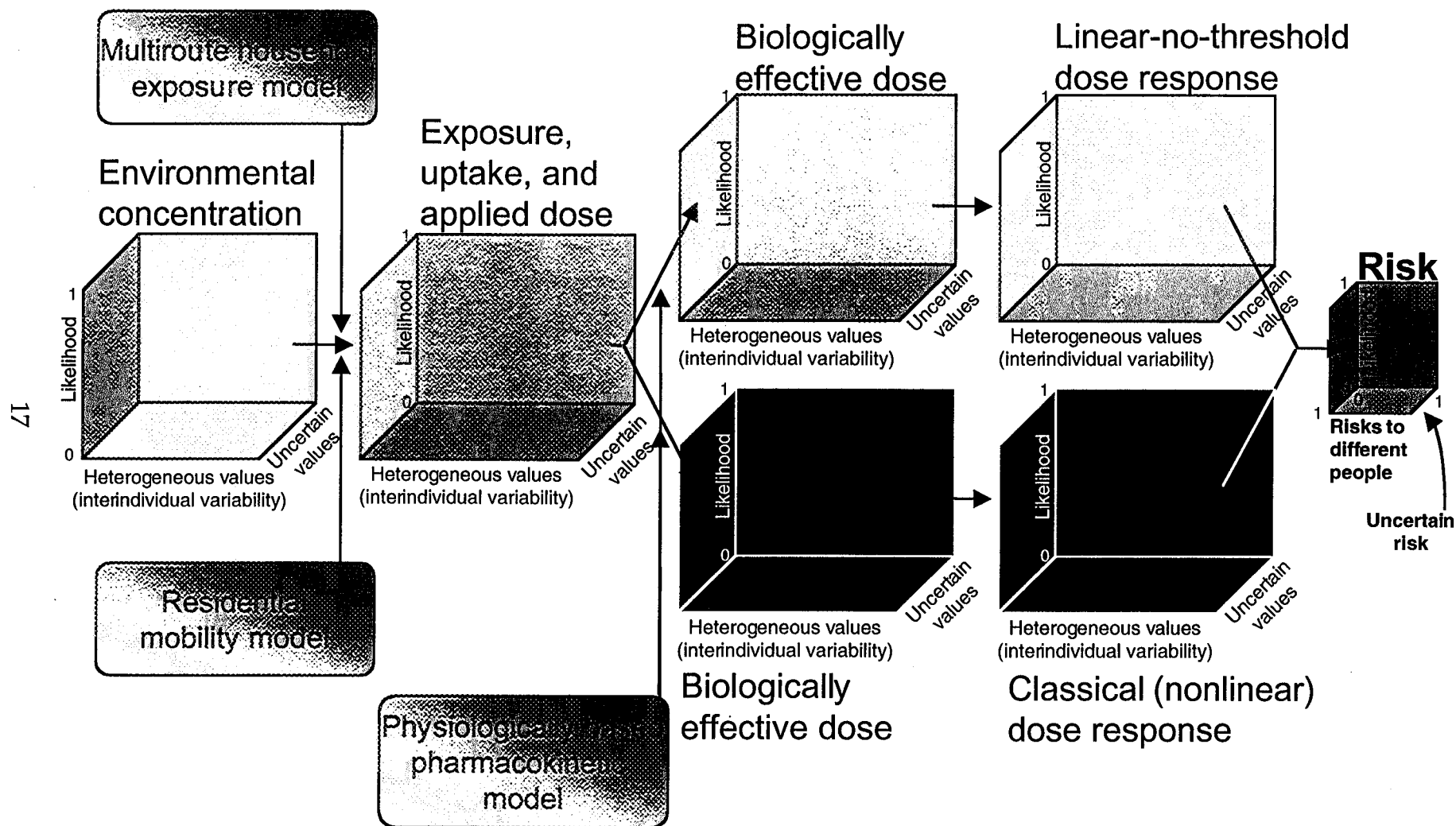
Figure 1. Unified probabilistic approach to risk assessment for TCE.

1999). (Note that the Phase 1 report used slightly different subscript-notation.) Specifically, it was assumed that,

$$X_{G,ing} = E_{Ing} \ , \tag{1a}$$

$$X_{G,der} = E_{Derm} \ , \tag{1b}$$

$$X_{G,inh} = (V_{alvR}/Inh) E_{Inh} \ , \quad \text{where} \tag{1c}$$

Inh = total respiratory ventilation rate used in Daniels et al. (2000) ($m^3 \ kg^{-1} \ d^{-1}$); and

$$V_{alvR} = \frac{V_{alvr}}{1000 \ L \ m^{-3}} V_{alv} = \frac{12.9 \ L \ h^{-1}}{1000 \ L \ m^{-3}} (V_W/kg)^{0.74-1} V_{alv} \ , \quad \text{in which:} \tag{2}$$

V_{alvR} = weight-normalized ventilation rate used in present study ($m^3 \ h^{-1}$),

V_{alvr} = non-normalized ventilation rate ($L \ h^{-1}$),

V_W = body weight (kg), and

V_{alv} = normalized interindividual variability in V_{alvR} that is independent of variability in V_W (unitless).

Equation 2 is an adaptation of the alveolar ventilation rate V_{alvr} defined in a validated PBPK model for TCE in humans (Allen and Fisher, 1993). Because, as explained in Section 2.3, this PBPK model was integrated into the probabilistic framework used in the present study, $X_{G,inh}$ was defined in terms of V_{alvR} rather than the total ventilation rate Inh used by Daniels et al. (2000).

Variability in V_W for U.S. adults was modeled as approximately lognormal (LN) with an arithmetic mean (AM) of 71.0 kg, standard deviation (SD) of 15.9 kg, and corresponding coefficient of variation (CV = SD/AM) of 0.224 (CalEPA, 1996; Finley et al., 1994). Based on the method of moments (Aitchison and Brown, 1957) explained and developed in Appendix 1, the assumed AM and SD of V_W imply that $V_W \sim LN(4.24, 0.221)$.

Weight-normalized rates of total respiratory ventilation for U.S. adults are approximately lognormally distributed with a CV of ~0.3 (CalEPA, 1996). As implied by Equation 1, the non-normalized alveolar ventilation rate V_{alvr} ($L \ h^{-1}$) is approximated as

$$V_{alvr} \approx (12.9 \ L \ h^{-1}) (V_W \ kg^{-1})^{-0.3} V_{alv} \ . \tag{3}$$

18

The alveolar proportion of total lung volume was assumed to be nearly constant, and consequently variability in V_{alv} was modeled as LN with an AM of 1. Based on the method of moments (Appendix 1), it follows that $V_{alv} \sim LN(-0.0409, 0.286)$. To facilitate PBPK analyses described in Section 2.3, it was assumed that V_{alv}, V_{alvr} and V_{alvR} pertain to children as well as to adults.

Predicted daily (non-LTWA) peak TCE exposures ($X_{C,P}$, in mg kg^{-1} d^{-1}) due to Site LF-13 groundwater contamination at Beale AFB were defined for ingestion and dermal exposure routes as follows, in terms of $X_{G,P}$ defined above (Equations 1a-b) and in terms of the constants EF and AT and the variate ED defined in Equation 1 of the Phase-1 report (Daniels et al., 2000):

$$X_{C,ing} = [AT/(ED \times EF)]X_{G,ing} , \quad \text{and} \tag{4a}$$

$$X_{C,der} = [AT/(ED \times EF)]X_{G,der} , \quad \text{where} \tag{4b}$$

ED = household exposure/residence duration used in Daniels et al. (2000) (y);

EF = exposure frequency used in Daniels et al. (2000) (d y^{-1});

AT = averaging time used in Daniels et al. (2000) corresponding to a 70-y exposure (d).

Predicted daily peak respiratory TCE exposure, $X_{C,inh}$, was similarly related to $X_{G,inh}$, except that E_{Inh} (see Equation 1c) was defined by Daniels et al. (2000) to refer to total household LTWA exposure, whereas $X_{C,inh}$ pertains to peak respiratory TCE exposure, which is assumed to occur during showering (and without reference to non-shower respiratory TCE exposures). Therefore, $X_{C,inh}$ was modeled as follows based on the method of Daniels et al. (2000) as adapted in Equation 1c, but solely with reference to shower-related TCE exposure:

$$X_{C,inh} = V_{alvr}\left(\frac{W_{sh}\ \phi_{TCE-sh}}{1000\ \text{L m}^{-3} \times AE_{sh}}\right)C_w\frac{V_{t,inh}}{1\ \text{d}} , \quad \text{where} \tag{4c}$$

W_{sh} = water-usage rate per person for shower (L h^{-1});

ϕ_{TCE-sh} = water-to-air transfer efficiency of TCE in the shower (unitless);

AE_{sh} = air-exchange rate in the shower or bath stall (m^3 h^{-1});

C_w = TCE concentration in ground water (mg L^{-1}); and

$$V_{t,\text{inh}} = (ET_{\text{sh}} \times 1 \text{ d}) = \text{shower duration (h)};$$

where variability in W_{sh}, $\phi_{\text{TCE-sh}}$, AE_{sh}, and ET_{sh}, and uncertainty in C_w, were all modeled as previously described (Daniels et al., 2000).

2.3. Biologically Effective Dose

For reasons discussed in Section 1.3, liver was assumed to model susceptible target tissue for TCE-induced cancer based on mouse bioassay data, and mouse hepatocellular toxicity was used to model the most sensitive TCE-induced noncancer (but possibly cancer-related) endpoint. Dose-response relations for TCE-induced endpoints were treated as functions of corresponding mechanism- and route-specific measures of biologically effective dose $D_{\text{MA,P}}$ (mg kg^{-1} d^{-1}) defined below. As indicated in Figure 1, PBPK and associated JUV models used to define $D_{\text{MA,P}}$ as functions of corresponding TCE exposures ($X_{\text{MA,P}}$) were treated differently in view of uncertainty as to the extent to which the MA for TCE involves genotoxic (G) processes with a plausibly linear dose-response vs. cytotoxic/mitogenic (C) processes with a likely nonlinear dose-response. To facilitate subsequent calculations, the following related quantities were also calculated:

$$D_{\text{MA}} = \sum_{i=\{\text{ing,inh,der}\}} D_{\text{MA},i} , \tag{5a}$$

$$f_{\text{MA,P}} = \left\langle \overline{D_{\text{MA,P}}} \right\rangle / \left\langle \overline{D_{\text{MA}}} \right\rangle , \tag{5b}$$

$$B_{\text{MA,P}} = D_{\text{MA,P}} / \left\langle \overline{D_{\text{MA,P}}} \right\rangle , \quad \text{and} \tag{5c}$$

$$B_{\text{MA}} = \sum_{i=\{\text{ing,inh,der}\}} \frac{B_{\text{MA},i}}{3} \left| \left(F_{B_{\text{MA},i}}(B_{\text{MA},i}) = p \right) \right. , \quad \text{for } 0 \leq p \leq 1, \quad \text{where} \tag{5d}$$

D_{MA} = total of all pathway-specific biologically effective doses under mechanistic assumption MA (mg kg^{-1} d^{-1} for MA = G, mg/L for MA = C);

$B_{\text{MA,P}}$ = normalized biologically effective dose for pathway P under mechanistic assumption MA (unitless);

B_{MA} = mean value of the cumulative probability distribution functions (cdfs) corresponding to all pathway-specific normalized biologically effective doses under mechanistic assumption MA (unitless); and

$f_{\text{MA,P}}$ = fraction of $\left\langle \overline{D_{\text{MA}}} \right\rangle$ due to pathway P (unitless).

The variate B_{MA} is defined (by Equation 5d) to take advantage of the fact that MA-specific $B_{MA,P}$ distributions were found to be nearly identical for all pathways P in this case study (see Results). Therefore, in calculations implementing Monte Carlo simulations done in this study, EB_{MA} distributions were used in place of the greater number of corresponding $ED_{MA,P}$ distributions (see Section 2.6). The order of (uncertainty- vs. variability) expectation operations in Equations 5a-b is arbitrary in this case study, because the order was not found to have a substantial effect on the value of $\langle \overline{D_{MA}} \rangle$ obtained (due principally to the linear structure and behavior of the models used for $D_{MA,P}$ previously described; see Daniels et al., 2000). Consequently, $\langle \overline{B_{MA}} \rangle = \overline{\langle B_{MA} \rangle} = 1$ by definition. Note also that, because of the multiplicative model structures implied by Equations 4a-c, $B_{MA} = \langle B_{MA} \rangle \overline{B_{MA}}$.

The rationale for including both genotoxic and cytotoxic MAs into this analysis is discussed below, followed by subsections detailing PBPK models and methods used to calculate corresponding biologically effective genotoxic and cytotoxic doses to bioassay animals and to humans.

2.3.1. Uncertainty in mechanism of toxic action

As discussed in Section 1.3, there is fundamental "model" uncertainty regarding critical mechanism(s) explaining the observed ability of TCE to increase tumor incidence in rodent bioassays and its suspected ability to do the same in humans. This uncertainty can be represented by the following four alternative mechanistic assumptions (MAs):

Assumption 1 (MA_G) is the traditional approach to assessing TCE cancer risk, which presumes that TCE increases cancer risk only via one or more genotoxic mechanisms of action, involving DNA damage that is linearly proportional to the biologically effective concentration of one or more of TCE's reactive metabolites (Bogen, 1988; Bogen et al., 1988; Brown et al., 1990; USEPA, 1985,1987a).

Assumption 2 (MA_C) is that observed TCE-induced (e.g., liver) cancer is due entirely to increased net proliferation of spontaneous premalignant cells elicited primarily by TCA, by a cytotoxic and/or perhaps a directly mitogenic mechanism (Andersen et al., 1998; Bogen and Gold, 1997).

Assumption 3 ($MA_{G \cap C}$) is the composite assumption that both genotoxic and nongenotoxic mechanisms contribute to observed TCE carcinogenicity in

21

bioassays, i.e., that both MA_G *and* MA_C are true. However, to the extent $MA_{G \cap C}$ is true, uncertainty remains as to the quantitative role played by each mechanism involved. This kind of uncertainty is often referred to as "parameter" uncertainty, because it is possible to reflect this as uncertainty pertaining to a single parameter (in a sufficiently general model) that governs the weight to be given to each of the two mechanisms considered to be operative under $MA_{G \cap C}$.

Assumption 4 ($MA_{G \cup C}$) is the "dichotomous" assumption that *either* MA_G *or* MA_C is true, but there is "model" uncertainty as to which one of these possibilities is true, in view of the fact that the "parameter" uncertainty discussed above in reference to $MA_{G \cap C}$ is quantitatively equivalent to "model" uncertainty.

In view of evidence discussed in Section 1.4 supporting the plausibility of both MA_G and MA_C, both of these mechanistic assumptions were used to define route-specific biologically effective dose and dose-response for TCE-induced cancer. Of course, MA_C was used exclusively as the basis for calculating biologically effective dose and dose-response for TCE-induced noncancer endpoints. Below, methods used to estimate biologically effective doses corresponding to mechanisms MA_G and MA_C are described, following an explanation of the PBPK modeling approach that was adopted in this study to accommodate both mechanisms.

2.3.2. PBPK modeling approach

A number of multi-compartment PBPK models have been developed that provide reasonably well-validated descriptions of the uptake, distribution, metabolism, and excretion of TCE administered by various routes to mice, rats and humans (Abbas and Fisher, 1997; Allen and Fisher, 1993; Bogen, 1988; Fisher and Allen, 1993; Fisher et al., 1991,1998; Stenner et al., 1998; USEPA, 1985, 1987b). In contrast to earlier PBPK models describing TCE distribution, metabolism and excretion using four physiological compartments, the more recent "second generation" models include additional compartments to describe distribution, metabolism and excretion of TCA and of unbound and glucuronide-bound trichloroethanol in mice and humans (Abbas and Fisher, 1997; Fisher et al., 1998), and to account for enterohepatic recirculation of TCA and of trichloroethanol-glucuronide (Stenner et al., 1998). Although the newer PBPK models are more realistic, they are less convenient to incorporate into the adopted probabilistic framework relative to earlier-type 4-compartment models. It is also not apparent that any improved ability to fit empirical data used to validate the newer vs.

22

the earlier models implies any corresponding substantial improvement in the specific measures of biologically effective dose discussed below, namely, total metabolized TCE and peak plasma concentration of TCA. Indeed, an earlier-type 4-compartment model for TCE in humans appears to provide fairly accurate predictions of the peak value, $Max(C_{TCA})$, of TCA concentration in plasma measured in several different studies involving humans exposed by inhalation to various air concentrations of TCE (Allen and Fisher, 1993), whereas a corresponding "second generation" model appears to underpredict $Max(C_{TCA})$ by up to ~40% in human subjects exposed to 50 or 100 ppm TCE in air (Figure 8 of Stenner et al., 1998). Therefore, earlier-type 4-compartment models (Allen and Fisher, 1993; Bogen, 1988) were used for PBPK-based calculations of biologically effective dose in the present study, as described below. However, recently reported experimental data on human variability in key PBPK parameter values (Fisher et al., 1998; Lipscomb et al., 1998) was incorporated into the present analysis as discussed below.

2.3.3. Effective genotoxic dose

Under MA_G for TCE (i.e., assuming that TCE is a "linear"/genotoxic carcinogen), bioassay-based potency traditionally has been expressed as increased risk per unit of PBPK-estimated total LTWA metabolized TCE per kg body weight per day, without accounting for PBPK-related uncertainty and variability (Bogen, 1988; Brown et al., 1990; USEPA, 1985, 1987b). There is an indication this policy will likely persist (USEPA, 1996). Measures of biologically effective dose, as LTWA metabolized TCE in mg kg^{-1} d^{-1} to animals in bioassays positive for TCE-induced liver or kidney cancer, were obtained from Table 4 of Bogen (1988). Similar measures of route-specific biologically effective dose $D_{G,P}$ to humans under MA_G were used for the present analysis, namely:

$$D_{G,P} = V_{fm,P} X_{G,P} , \quad \text{for P = ing, inh, or der, where} \tag{6}$$

$V_{fm,P}$ = limiting fraction of total TCE intake by pathway P that is metabolized conditional on intake being sufficiently small to ensure that saturation of TCE metabolism remains negligible (unitless).

For multi-compartment PBPK models like that of Allen and Fischer (1993), these limiting metabolized fractions were shown previously to be

$$V_{\text{fm,ing}} = \left[1 + \frac{K_\text{m}}{V_\text{max}}\left(\frac{V_\text{Pb}}{V_\text{alvr}} + \frac{1}{V_\text{liv}}\right)^{-1}\right]^{-1} \quad , \text{ and} \tag{7a}$$

$$V_{\text{fm,inh}} = V_{\text{fm,der}} = \left[1 + \frac{V_\text{alvr}}{V_\text{Pb}}\left(\frac{K_\text{m}}{V_\text{max}} + \frac{1}{V_\text{liv}}\right)\right]^{-1} \quad , \text{ where} \tag{7b}$$

V_alvr = alveolar ventilation rate (defined in Equations 2 and 3) (L h^{-1});

V_liv = the rate of blood perfusion to liver (L h^{-1});

V_Pb = the blood:air partition coefficient for TCE (L$_\text{air}$ L$_\text{blood}^{-1}$);

V_max = maximum rate of TCE metabolism (mg h^{-1});

K_m = Michaelis-Menten affinity/saturation constant (mg L^{-1}); and

where V_alvr is alveolar ventilation rate defined above (Equation 3), and where, for Michaelis-Menten parameters K_m and V_max assumed to govern metabolic saturation kinetics for TCE, the mass unit (mg) refers to TCE and the volume (L) to venous blood exiting liver (Bogen, 1988; Bogen and Hall, 1989).

To derive human biologically effective doses under MA$_\text{G}$ (as well as under MA$_\text{C}$, as explained below), Equations 7a-b were applied assuming: V_alvr is defined by Equation 2, $V_\text{liv} = 26\% \times (15.0/12.9) \times V_\text{alv}$ (Allen and Fisher, 1993), $K_\text{m} = 1.5$ mg L^{-1} (i.e., treated as a constant) (Allen and Fisher, 1993), and that $V_\text{Pb} \sim N(10.2, 1.6)$ for males and females combined (Fisher et al., 1998). Variability in the maximal rate of TCE metabolism, V_max, was modeled as LN with

$$V_\text{max} = (14.9 \text{ mg h}^{-1})(V_\text{W} \text{ kg}^{-1})^{-0.3} V_\text{Vmax} \quad , \tag{8}$$

which adapts the definition used by Allen and Fisher (1993) to incorporate a multiplicative factor V_Vmax reflecting V_max-related variability, where V_Vmax was assumed to have an AM of 1. Under these assumptions, Equations 7a-b are simplified to:

$$V_{\text{fm,ing}} = \left\{1 + V_\text{alv}\left[V_\text{Vmax}(0.7700\,V_\text{Pb} + 2.547)\right]^{-1}\right\}^{-1} \quad , \text{ and} \tag{9a}$$

$$V_{\text{fm,inh}} = V_{\text{fm,der}} = \left[1 + \frac{V_\text{alv}}{V_\text{Pb}}\left(\frac{1.299}{V_\text{Vmax}} + 3.307\right)\right]^{-1} \quad , \tag{9b}$$

24

in which no more than three significant figures are implied. From Equation 9b it is clear that $V_{fm,inh}$ is correlated with V_{alv}. From Equation 1c, it follows that this correlation is implied in Equation 6 defining $D_{G,inh}$, as well as in Equation 15b below (in Section 2.3.4) that defines the corresponding cytotoxic dose $D_{C,inh}$. Note, however, that the limiting metabolized fractions defined by Equations 9a-b are *independent* of body weight (V_W), and thus are independent of $D_{MA,P}$ for P = {ing, der} defined by Equation 6 (and by Equations 15a-b below).

Based on *in vitro* measures of V_{max} for TCE using human microsomes and hepatocytes sampled from 4 to 6 different donors (Lipscomb et al., 1998), the CV of V_{max} was estimated to be approximately 0.60. Based on the method of moments (Appendix 1), and conditional on assumed variability in V_W discussed above (after Equation 2), the latter CV estimate implies that $V_{max} \sim LN(-0.152, 0.551)$. Systematic uncertainties pertaining to $V_{fm,P}$ are likely to be small relative to the combined effect of interindividual pharmacokinetic variabilities, so uncertainty *per se* is not incorporated into Equations 9a-b used to define $V_{fm,P}$.

Note that, conditional on the adopted PBPK model, Equations 7a-b and 9a-b remain true *regardless* of any (dynamic or static) pattern of exposure(s) involved, provided that metabolism remains virtually unsaturated, which in turn ensures that the corresponding system of linked ordinary differential equations remains linear (Bogen, 1988; Bogen and Spear, 1987).

2.3.4. Effective cytotoxic dose

Under MA_C for TCE, hepatocellular oxidative damage is assumed to comprise or elicit premalignant liver-cell proliferation and consequent increased tumor risk in mice, and is further assumed to correlate best with the daily peak value $Max(C_{TCA})$, of TCA concentration in plasma, rather than with LTWA total metabolized TCE or related areas under concentration-times-time curves for blood or other tissues (Bogen and Gold, 1997). Similar reliance on peak rather than LTWA metabolic yield was used for MA_C-based risk assessment for chlorinated methanes, based on empirical evidence supporting the former measure as the best predictor of oxidative damage (Bogen, 1990a). In the absence of dose-response data on TCA-induced rodent nephrotoxicity, and consistent with information discussed in Section 2.4, $Max(C_{TCA})$ was also taken to be

25

the biologically effective cytotoxic dose for potential TCE-induced kidney cancer under MA_C. It was further assumed that $Max(C_{TCA})$ is the biologically effective cytotoxic dose for TCE-induced noncancer endpoints, so in general it was assumed that $D_{C,P}$ for any exposure pathway P is the value of $Max(C_{TCA})$ produced in response to a corresponding exposure $X_{C,P}$ defined in Equation 4a-c. Corresponding total effective dose (D_C) was (in Equation 5a) defined as the sum of $D_{C,P}$ from all exposure pathways, as discussed below following Equation 13b.

In the context of low-dose risk extrapolation for humans based on the PBPK model for TCE used here, all saturable (Michaelis-Menten) PBPK relations linearize. Therefore, this PBPK model was evaluated using an entirely analytic approach previously described (Bogen and Gold, 1997), which is simpler yet equivalent to alternative, relatively cumbersome numerical methods more commonly applied. By this approach (see Equation 4 of Bogen and Gold, 1997),

$$\frac{dC_{TCA}(t)}{dt} = \frac{K_{fTCA}\,K_{MW}}{V_W V_{fd}}\left(\frac{C_{TCE}(t)\,V_{max}}{C_{TCE}(t)+K_m}\right) - V_{ke}\,C_{TCA}(t) \quad , \quad \text{where} \tag{10}$$

$C_{TCA}(t)=$ concentration at time t of TCA in plasma (mg L^{-1});

$C_{TCE}(t)=$ concentration at time t of TCE in venous blood exiting liver (mg L^{-1});

K_{fTCA} = net effective fraction of total TCE intake metabolized to TCA (unitless);

K_{MW} = TCA to TCE molecular-weight ratio (unitless); and

V_{fd} = fraction of body weight corresponding to apparent volume of distribution for TCA (L kg^{-1}); and

V_{ke} = first-order rate constant for elimination of TCA from plasma (h^{-1});

and where V_W, V_{max}, and K_m were defined above (after Equations 2, 7b, and 7b, respectively). It was assumed that K_{fTCA} = 0.33 (Allen and Fisher, 1993), and the ratio K_{MW} is 1.228 (see Bogen and Gold, 1997).

Conditional on any regular pattern of peak daily TCE exposures $X_{C,P}$ that—by any pathway P and corresponding duration $V_{t,P}$—are small enough to ensure that $C_{TCE}(t) <<$ K_m for all t, Equation 10 implies that $C_{TCA}(t)$ attains a dynamic equilibrium in which

$$Max[C_{TCA}(t)] = D_{C,P} = \frac{K_{fTCA}\,K_{MW}\,V_{fm,P}}{V_{fd}\,V_{ke}}\left(X_{C,P}\times 1\,d\right)\left[\frac{f_{deq}}{V_{t,P}}\right] \quad , \quad \text{where} \tag{11}$$

$$f_{deq} = \frac{1 - \exp(-V_{ke} V_{t,P})}{1 - \exp(-V_{ke} \times 24\,h)} \quad , \tag{12}$$

$V_{t,P}$ = duration ($\leq 24\,h$) of peak daily exposure $X_{C,P}$ by pathway P (h); and

f_{deq} = fraction of $Max[C_{TCA}(t)]$ conditional on a hypothetical infinite exposure duration that is attained at dynamic equilibrium conditional on $V_{t,P}$;

in which $X_{C,P}$ and $V_{t,inh}$ were defined via Equations 4a-c, $V_{fm,P}$ was defined in Equations 9a-b, and all other variates in Equations 11-12 (i.e., besides f_{deq}, $V_{t,P}$, and $X_{C,P}$) were defined following Equation 10. Equation 11 is a multi-route generalization of Equations 6 and 7 of Bogen and Gold (1997). Figure 2 shows how the bracketed term in Equation 11 is well approximated by

$$\left[f_{deq} / V_{t,P} \right] \approx (24\,h)^{-1} + 0.5053\,V_{ke} + (1.661\,h)V_{ke}^2$$
$$\text{for } V_{t,P} \leq 0.5\,h \text{ and } V_{ke} \leq 0.1\,h^{-1} \tag{13a}$$

$$\approx (24\,h)^{-1} + 0.5053\,V_{ke} \qquad \text{for } V_{t,P} \leq 0.5\,h \text{ and } V_{ke} \leq 0.04\,h^{-1} \; , \tag{13b}$$

in which no more than three significant figures are implied.

As indicated following Equation 4c, variability in $V_{t,inh}$ was modeled as LN with $V_{t,inh} \sim LN(\ln 0.120, \ln 1.47)$ as previously described (Daniels et al., 2000), implying shower (or, more generally, bathing-related water-flow) durations that virtually never ($p < 10^{-11}$) exceed 0.5 h. It was further assumed that $V_{t,der} = V_{t,inh}$ as previously described (Daniels et al., 2000), that $V_{t,ing} < 0.5\,h$, and that $V_{t,ing}$, $V_{t,inh}$, and $V_{t,der}$ are timed such that the total effective cytotoxic exposure, D_C (defined by Equation 5a), is maximized, so as to reflect the peak value of $Max[C_{TCA}(t)]$ predicted during a lifetime of different pathway-specific effective-exposure scenarios.

Based on methods used and data reported by Allen and Fisher (1993), variability in V_{fd} was modeled as uniformly distributed between 5.2 and 15.2%, and as being negatively correlated with V_W with a rank correlation coefficient of $\rho_r(V_{fd}, V_W) = -0.50$; and it was assumed that

$$V_{ke} = 0.028\,V_W^{-0.3}\,V_e \; , \quad \text{where} \tag{14}$$

V_e = normalized variability in V_{ke} that is independent of variability in V_W (unitless),
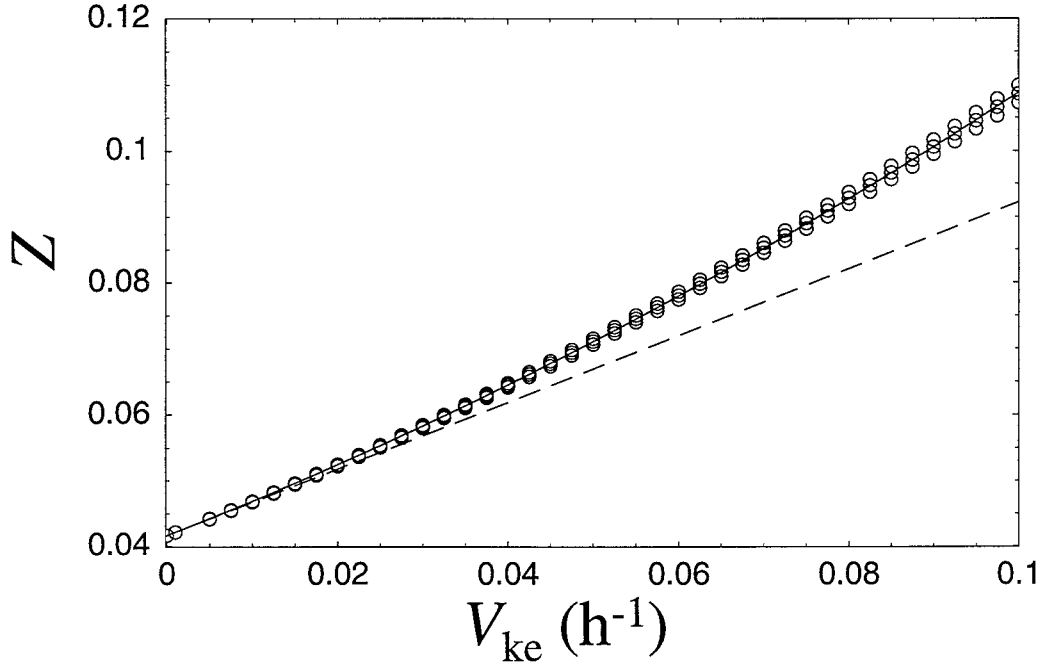
Figure 2. Approximation of the ratio $Z = (f_{deq}/V_{t,P})$, i.e., the fraction of steady-state that is attained under dynamic-equilibrium exposure conditions, divided by the duration $V_{t,P}$ of daily exposure pulses. Z (unitless) is plotted (using open points) as a function of 40 values of the (heterogeneous) TCA-elimination rate, V_{ke}, evenly spaced between 0 and 0.1 h^{-1}. The relatively small amount of vertical variation in the plotted points corresponds to three different values of $V_{t,P}$ used (0.01, 0.25, and 0.5 h) conditional on each value of V_{ke} used. To these points was fitted the linear quadratic curve shown: $Z = (0.05053\ h)V_{ke} + (1.6608\ h^2)V_{ke}^2$. For $V_{ke} < 0.04\ h^{-1}$, the relative error of this fit using only the linear term (0.05053 h) is <5%.

28

and where V_W was previously defined (after Equation 2). It was further assumed that variability in V_{ke} is lognormally distributed. Experimental data reported for 17 male and female human subjects indicates that V_{ke} has a CV of ~0.60 (Fisher et al., 1998). Based on the method of moments (Appendix 1; see discussion concerning V_{max} following Equation 8), it follows that V_e has an AM of 1 and that $V_e \sim$ LN(-0.152, 0.551). Consequently, Equation 14 implies that $V_{ke} < 0.030$ h^{-1} for virtually (>99% of) all modeled individuals at risk. Because $V_{t,ing} < 0.5$ h is assumed as described above, Approximation 13b is accurate (to within <2.5%), and was thus used to evaluate Equation 11. These two equations, together with assumptions stated above, yield:

$$D_{C,P} = \left(X_{C,P} \times 1\,\mathrm{d}\right)\frac{V_{fm,P}}{V_{fd}}\left(\frac{0.6107\,V_W^{0.3}}{V_e} + 0.2074\right) \;,\; \text{for P} = \{\text{ing, der}\}, \text{ and} \qquad (15a)$$

$$D_{C,inh} = \left(X_{C,inh} \times 1\,\mathrm{d}\right)\frac{V_{fm,inh}}{V_{fd}}\left(\frac{7.878}{V_e} + \frac{2.674}{V_W^{0.3}}\right) \;, \qquad (15b)$$

in which no more than three significant digits are implied, and where: $X_{C,P}$ for pathways P = {ing, inh, or der} were defined by Equations 1a-c, V_W was defined after Equation 2, $V_{fm,P}$ for pathways P were defined in Equations 9a-b, V_{fd} was defined after Equation 10, and V_e was defined after Equation 14. Note that V_{fd} and V_W are assumed to be correlated (as discussed prior to Equation 14), as are V_{alv} and V_{inh} as discussed above (in Section 2.3.3, after Equation 9b).

2.4. TCE Dose-Response

The following subsections discuss methods used to model dose-response for TCE-induced cancer and noncancer endpoints, and associated JUV. Sections 2.4.1 and 2.4.2 describe methods used for dose-response modeling under MA$_G$ and MA$_C$, respectively. Section 2.4.3 then describes the method used to incorporate uncertainty concerning the mechanism of action for TCE-induced cancer into estimates of cancer risk as well as of corresponding aggregate (cancer and noncancer) risk.

2.4.1. Dose-response assuming genotoxic mechanism(s)

Under MA$_G$, linear-no-threshold extrapolation of TCE cancer risk is based on the assumption that TCE can increase cancer risk via one or more genotoxic mechanisms of

action. These mechanisms involve DNA damage that is presumed to be linearly proportional to the biologically effective concentration of one or more of TCE's reactive metabolites, where potency is estimated for each bioassay in terms of a pharmacologically based equivalent effective dose—namely, the total amount of TCE metabolized per kg body weight per day (Bogen, 1988; USEPA, 1985, 1987b). Effective bioassay doses $D_{C,P}$ and corresponding positive, malignant (plus, where applicable, benign) tumor responses in mouse liver and rat kidney were obtained from information listed in Table 4 of Bogen (1988) concerning seven rodent bioassay data sets (Bell et al., 1978; Maltoni et al., 1986; NCI, 1976; NTP, 1990). The studies involved are summarized below in Table 1. For each data set, a cdf reflecting uncertainty (estimation-error) in estimated cancer potency (i.e., "slope factor", or risk per unit dose), here denoted U_{pot} (kg d mg^{-1}), was calculated as described below (Section 2.6).

A subjective weighting scheme was then used to address uncertainty associated with lack of knowledge concerning which of the multiple positive animal bioassay results for TCE in rodent liver and kidney best predicts TCE cancer risk in humans, similar to an approach previously applied to characterize JUV in cancer risk posed by environmental exposure to chloroform (Bogen, 1995). To each species/sex-specific potency distribution obtained as described above, the corresponding relative weight indicated in Table 1 was applied to obtain a single weighted-average distribution reflecting uncertainty in tumor likelihood, conditional on effective dose. (This weighted average was obtained analytically, via calculations analogous to those indicated in Equation 5d.) The weights used assign equal likelihood (of reflecting true carcinogenic potency in humans) to bioassay data sets that differ: by sex within a given strain, by strain within a given species, and by species.

Animal-to-human extrapolation of toxicokinetically equivalent effective dose was done by using an appropriate PBPK model as described above, so no additional factor was employed in this regard in accordance with currently proposed policy (see Section 1.4). An uncertain factor U_{tdyn} was used to account for interspecies toxicodynamic dynamic differences between rodents and humans (i.e., in increased likelihood of cancer per unit effective genotoxic dose). Analogous to toxicodynamic factors recommended recently by the USEPA for noncarcinogens and "nonlinear"

Table 1. Bioassay data sets used to estimate potency of TCE as a genotoxic/linear liver or kidney carcinogen

No.	Study[a]	Species	Strain	Sex	Route[b]	Tumor type[b]	No. dose grps.	Relative study weight[c]
1	NCI (1976)	mouse	B6C3F1	M	gav	HCC	3	1
2	NCI (1976)	mouse	B6C3F1	F	gav	HCC	3	1.5
3	NTP (1990)	mouse	B6C3F1	M	gav	HCA	2	1
4	NTP (1990)	mouse	B6C3F1	F	gav	HCA	2	1.5
5	NTP (1990)	rat	F344	M	gav	RTCA	3	12
6	Bell et al. (1978)	mouse	B6C3F1	M	inh	HCA	3	1
7	Maltoni et al. (1986)	mouse	Swiss	M	inh	MH	3	6

[a]More detailed study-specific information appears in Table 4 of Bogen (1988).

[b]Lifetime bioassay exposure scenarios: gav = gavage 5 d/wk in oil vehicle; inh = inhalation 6 h/d, 5 d/wk. Tumor types: HCC = hepatocellular carcinomas; HCA = HCC or hepatocellular adenomas; RTCA = renal tubule-cell carcinomas or adenomas; MH = malignant hepatomas.

[c]Assigned *a priori* relative study weight (see text).

carcinogens (USEPA, 1998), but using a probabilistic approach as previously proposed for noncancer endpoints (Carlson-Lynch et al., 1999; Slob and Pieters, 1998), it was assumed U_{tdyn} is lognormally distributed, has a GM of 1 (i.e., is as likely as not to exceed 1), and is unlikely (p < 0.01) to exceed a value of 3. A similar factor V_{tdyn} used to reflect intraspecies toxicodynamic variation was assumed to have an AM of 1 and to be unlikely (p < 0.01) to exceed 10. By the method of moments (Appendix 1), it thus was assumed that $U_{tdyn} \sim LN(0, 1.60)$ and $V_{tdyn} \sim LN(0.700, 2.33)$.

Combining the dose-response factors discussed above, and noting that $\overline{V_{tdyn}} = 1$, increased risk R_G under MA_G was defined using a low-dose-linear multistage risk-extrapolation model as

$$R_G = 1 - \exp\left\{-\left[U_{pot}U_{tdyn}V_{tdyn}\langle\overline{D_G}\rangle \sum_{P=\{ing,inh,der\}} f_{G,P}B_{G,P}\right]\right\} , \qquad \text{with} \quad (16a)$$

$$\overline{R_G} \approx 1 - \exp\left(-U_{pot}U_{tdyn}\langle\overline{D_G}\rangle\,\overline{B_G}\right) , \qquad \text{and} \quad (16b)$$

$$\langle R_G\rangle \approx 1 - \exp\left\{-\left[\langle U_{pot}\rangle\langle U_{tdyn}\rangle V_{tdyn}\langle\overline{D_G}\rangle \sum_{P=\{ing,inh,der\}} f_{G,P}\langle B_{G,P}\rangle\right]\right\} , \qquad (16c)$$

in which U_{pot}, U_{tdyn}, and V_{tdyn} were defined above in this subsection, and the remaining variates were defined in/after Equations 5a-b with reference to Equation 6. In Approximation 16b, $\overline{B_G} = \overline{B_{G,P}}$ because (conditional on Equations 1a-c, on Equation 6, and on all heterogeneous variates involved in $B_{C,P}$), uncertainty in $B_{C,P}$ is due entirely to uncertainty in the variates ED and C_w (defined after Equations 4b and 4c, respectively) that are both independent of pathway P. Note that the $\langle B_{G,P} \rangle$ variates in Approximation 16c are correlated (see Section 2.6). Approximations 16b-c are 1st-order approximations (see Bogen and Spear, 1987). However, for extrapolation of risks $<10^{-2}$ the functions involved are effectively linear, so the approximations entail only negligible loss of accuracy.

2.4.2. Dose-response assuming cytotoxic mechanism(s)

The ability of TCE to induce cancer under MA_C was assumed to arise from TCA-induced cytotoxicity/mitogenicity indicated by increased formation of thiobarbituric-acid-reactive substances (TBARS), as previously suggested (Bogen and Gold, 1997). Absent better data, increased TBARS elevation above background was modeled using data on male B6C3F1 mice administered a single gavage dose of 0, 100, 300, 1000, or 2000 mg TCA per kg body weight in buffered water, and corresponding measured peak TCA concentrations in plasma, $Max(C_{TCA})$ (Larson and Bull, 1992a). Multiple independent interactions are likely to be involved in TCA-induced oxidative-stress. Consequently (Aitchison and Brown, 1957), dose-response under MA_C could reasonably be modeled using a two-parameter LN function

$$Y(A_C) - Y_0 = 100\,\Phi\left(\frac{\log_{10}(A_C) - \mu}{\sigma}\right) , \quad \text{in which} \tag{17}$$

A_C = administered acute TCA dose (mg TCA per kg body weight);

$Y(A_C)$= TBARS level (nmol malondialdehyde equiv. per g liver) induced by A_C;

Y_0 = $Y(0)$ = background TBARS level (nmol malondialdehyde equiv. per g liver);

$\Phi(z)$ = cumulative probability distribution function (cdf) of a standard normal (Gaussian) random variate Z, equal to $\text{Prob}(Z \leq z)$;

μ = location parameter = $\log_{10}GM$ (unitless); and

σ = shape parameter = $\log_{10}GSD$ (unitless);

to the mouse TBARS-vs.-A_C data (see Section 2.6), where the unit of $Y(A_C)$ is henceforth suppressed for convenience. The arbitrary constant (100) in the two-parameter model (Equation 16) was used because a three-parameter LN model fit to these data did not yield plausibly unique parameter estimates.

Raw $Y(A_C)$-vs.-A_C data ([4 measures]×[4 noncontrol dose levels]) summarized by Larson and Bull (1992a) were assumed to be approximately normally distributed, with a reported arithmetic mean (AM) and standard deviation (SD) for Y_0 of 40.0 and 4.0 nmol malondialdehyde equiv. per g liver, respectively. Error in this AM was modeled as t-distributed with 3 degrees of freedom. TBARS elevations above the corresponding 2-tail upper 95% confidence limit on Y_0 were assumed to be biologically significant in the sense of being plausibly related to TCA-induced cytotoxicity. This upper bound on Y_0 shall be denoted Y_{sig}, and the mouse data indicate that $Y_{sig} = 49.4$ nmol malondialdehyde equiv. per g liver. Estimates $\{\hat{\mu}, \hat{\sigma}\}$ of the LN parameters $\{\mu, \sigma\}$, as well as their corresponding estimated SDs $\{s_{\hat{\mu}}, s_{\hat{\sigma}}\}$ and product-moment correlation (r), were obtained by fitting the LN-model (Equation 17) to the mouse TBARS data, as described below (Section 2.6).

Parameter-estimation error pertaining to the likelihood that any particular acute administered dose A_C would induce a significantly elevated (i.e., presumptively cytotoxic) response was characterized by simulating uncertainty in parameters μ and σ. This was done by modeling these parameters as the corresponding variates $\mu = \hat{\mu} + s_{\hat{\mu}}U_{t1}$, and $\sigma = \hat{\sigma} + s_{\hat{\sigma}}U_{t2}$, where U_{t1} and U_{t2} are correlated uncertain variates each distributed as Student's T with $[(4 \times 4) - 2] = 14$ degrees of freedom, having a correlation coefficient equal to r. At the lowest A_C level used (100 mg kg^{-1}) in B6C3F1 mice, it was found that effective dose, expressed as the corresponding maximum plasma TCA concentration (D_{Ca}), was $D_{Ca} = Max(C_{TCA}) = 130$ mg L^{-1} (Larson and Bull, 1992a). For convenience, and absent data at lower D_{Ca} levels, it was assumed that the ratio D_{Ca}/A_C is a constant independent of A_C, equal to [(130 mg L^{-1})/(100 mg kg^{-1})] = 1.3 kg L^{-1}. Critical effective dose conditional on $\{\mu, \sigma\}$ was therefore modeled as

$$D_{Ca} = (1.3 \text{ mg L}^{-1}) 10^{\mu + \sigma \Phi^{-1}\left[(Y_{sig} - Y_0)/100\right]}$$

$$= (1.3 \text{ mg L}^{-1}) 10^{\left(\hat{\mu} + s_{\hat{\mu}} U_{t1}\right) + \left(\hat{\sigma} + s_{\hat{\sigma}} U_{t2}\right) \Phi^{-1}\left[(Y_{sig} - Y_0)/100\right]} \sim F_C(D_{Ca}) \text{ , where} \quad (18)$$

D_{Ca} = acute effective TCA dose (mg TCA per L plasma);

Y_{sig} = significantly elevated value of $Y(A_C)$ above Y_0 (defined above);

U_{ti} = (for i=1,2) correlated errors distributed as Student's T cdfs with df = 14; and

$F_C(D_{Ca})$ = cdf of D_{Ca}, specifying the modeled likelihood of significant TBARS elevation conditional on effective dose D_{Ca};

where Y_0, $\hat{\mu}$, $s_{\hat{\mu}}$, $\hat{\sigma}$, and $s_{\hat{\sigma}}$, as well as the t-distributed variates U_{t1} and U_{t2}, were all defined above, and where "~" means "is distributed as". Absent dose-response data on TCA-induced lipoperoxidation or cytotoxicity in rat kidney, it was assumed that $F_C(D_{Ca})$ also applies to rat kidney. This assumption is probably conservative, because relation between A_C and $D_{Ca} = \text{Max}(C_{TCA})$ was observed to be similar in rats vs. mice, whereas TCA is less effective at inducing TBARS elevation in rats vs. mice (Larson and Bull, 1992a).

Detailed dose-response information relating chronically or subchronically administered TCA and induced TBARS or cytotoxicity are still unavailable. Therefore, extrapolation of D_{Ca} to equivalent subchronic effective TCA dose, and extrapolation of subchronic to chronic effective TCA dose (where the latter is denoted D_C, and is defined by Equations 5a, 11, and 15a-b), was accomplished using the two uncertainty factors, U_{acute} and $U_{subchron}$, respectively. As previously suggested (Slob and Pieters, 1998), these factors were assumed to be lognormally distributed. Based on the observation that ratio of lowest observed effect levels for TCE-induced lethality in B6C3F1 mice is between 2 and 3 (NCI, 1976), it was assumed that U_{acute} has a GM of $GM_{acute} = 3$ and is unlikely ($p < 0.01$) to be greater than 6. The factor $U_{subchron}$ was assumed by default (see Slob and Pieters, 1998) to have a GM of $GM_{subchron} = \sim2$ and to be unlikely ($p < 0.01$) to exceed 10. Because many repeated daily exposures to TCE and/or its metabolite TCA are expected to be always more (never less) toxic than fewer exposures, it was assumed that a combined uncertainty factor (U_{chron}) extrapolates effective dose from a chronic to a toxicologically equivalent acute (bioassay) exposure condition as follows:

$$D_{Ca} = (1 + U_{chron})D_C \, , \quad \text{where} \tag{19a}$$

$$U_{chron} = U_{acute} \times U_{subchron}\left(1 - (GM_{acute}GM_{subchron})^{-1}\right) . \tag{19b}$$

That is, the combined LN factor U_{chron} was assumed to have a GSD equal to that of $U_{acute} \times U_{subchron}$ and a GM equal to one less than that of $U_{acute} \times U_{subchron}$. By the method of moments (Appendix 1), it was thus assumed that $U_{chron} \sim LN(\ln 5, \ln 2.12)$.

Animal-to-human extrapolation of toxicokinetically equivalent effective dose was done by using an appropriate PBPK model as described above, so no additional factor was employed in this regard in accordance with currently proposed policy (see Section 1.4). An uncertain factor U_{tdyn} was used to account for interspecies toxicodynamic dynamic differences between rodents and humans, and a similar factor V_{tdyn} to reflect intraspecies toxicodynamic variation, where U_{tdyn} and V_{tdyn} were defined above (in Section 2.4.1, prior to Equation 16).

Combining the dose-response factors discussed above, , and again noting that $\overline{V_{tdyn}} = 1$, increased risk R_C under MA_C was modeled as

$$R_C = F_C\left(U_{tdyn}V_{tdyn}(1+U_{chron})\langle\overline{D_C}\rangle \sum_{P=\{ing,inh,der\}} f_{C,P}B_{C,P}\right) , \quad \text{with} \tag{20a}$$

$$\overline{R_{C1}} = F_C\left(U_{tdyn}(1+U_{chron})\langle\overline{D_C}\rangle\,\overline{B_C}\right) , \quad \text{and} \tag{20b}$$

$$\langle R_C\rangle_1 = F_C\left(\langle U_{tdyn}\rangle V_{tdyn}(1+\langle U_{chron}\rangle)\langle\overline{D_C}\rangle \sum_{P=\{ing,inh,der\}} f_{C,P}\langle B_{C,P}\rangle\right) , \tag{20c}$$

in which U_{chron}, U_{tdyn}, and V_{tdyn} were defined above in this subsection; the remaining variates were defined in/after Equations 5a-c with reference to Equations 11 and 15a-b. In Approximation 20b, $\overline{B_C} = \overline{B_{C,P}}$ because (conditional on Equations 1a-c and 15a-b, and on all heterogeneous variates involved in $B_{C,P}$) uncertainty in $B_{C,P}$ is due entirely to uncertainty in the variate C_w (defined after Equation 4c), which in turn is independent of pathway P. Note that the $\langle B_{C,P}\rangle$ variates in Approximation 20c are correlated (see Section 2.6).

The subscripts "1" on the left side of Equations 20b-c each denote a 1st-order approximation, as was the case in Approximations 16b-c. However, Equations 20b-c

are expected to underestimate risk, because (in contrast to the exponentiated polynomial in Approximations 16b-c) F_C is a substantially nonlinear increasing function of effective dose (see Bogen and Spear, 1987). In the risk range from 10^{-10} to 10^{-2} relevant to this analysis, $\log(F_C(D_{Ca}))$ turns out to be well modeled by a linear-quadratic function of effective dose D_{Ca}, as explained below (Sections 2.6 and 3.2). Therefore, more accurate, corresponding 2nd-order approximations were used to evaluate R_C expectations, which were calculated as follows. For any risk R that is a function of a vector of n uncertain and/or heterogeneous variates $\mathbf{X} = \{X_1, X_2, ..., X_n\}$, the 2nd-order approximation of the expectation $ER(\mathbf{X})$—with respect only to variability, or only to uncertainty—of R, is given by (cf. Bogen and Spear, 1987):

$$E_2 R(\mathbf{X}) = E_1 R(\mathbf{X}) + \frac{1}{2}\sum_{i=1}^{k}\frac{\partial^2 R(\mathbf{X})}{\partial X_i^2}\sigma_{X_i}^2 \quad , \tag{21}$$

in which subscripts "1" and "2" on E denote the order of approximate expectation, $0 \leq k \leq n$, k = the size of the subset of $\mathbf{X}$ comprising all members of $\mathbf{X}$ that are variates of the type corresponding to expectation E (e.g., comprising all uncertain variates among $\mathbf{X}$ if E represents expectation with respect only to uncertainty in $R(\mathbf{X})$), σ^2 denotes the variance of the σ-subscripted variate with respect to its uncertainty or to its variability (again corresponding to how E is defined), and the second derivatives are all evaluated conditional on $\mathbf{X} = E\mathbf{X}$. Using Equation 21, second-order approximations were obtained to estimate conditional risks under MA_C corresponding to Equations 20b-c (see Section 2.6).

2.4.3. Model Uncertainty

In view of the plausibility of both MA_G and MA_C (see Section 1.3), uncertainty in the mechanism(s) of carcinogenic action for TCE was treated quantitatively, based on the "dichotomous" mechanistic assumption ($MA_{G \cup C}$) involving both MA_G and MA_C discussed above (Section 2.3). The alternative corresponding "composite" assumption ($MA_{G \cap C}$), which also involves both MA_G and MA_C (as discussed in Section 2.3), is more difficult to implement quantitatively than $MA_{G \cup C}$. $MA_{G \cap C}$ is more difficult to implement because it requires a complete model structure accounting for possible but unknown interactions between the different mechanisms considered. In contrast, $MA_{G \cup C}$ may be implemented simply by assigning the component assumptions (MA_G and MA_C)

corresponding, complementary *a priori* probabilities, and using the combination of these probabilities to reflect the (quantitatively equivalent) possibility that both MA_G and MA_C are true but to unknown degrees. $MA_{G\cup C}$ was therefore adopted using subjective probabilities U_G and $U_C = (1-U_G)$ to reflect the corresponding likelihoods that MA_G and MA_C, respectively, reflect the "true" mechanism of TCE-induced carcinogenic action. Consistent with our considered opinion that MA_C is *at least* as likely as not to explain observed TCE-induced cancer in rodent bioassays (see also Bogen and Gold, 1997), U_G was modeled as uniformly distributed between 0 and 0.5. Therefore, using de Morgan's rule (see Appendix I in NRC, 1994), increased *aggregate* risk R of incurring *either* genotoxicity-induced cancer, *or* cytotoxicity (which, in the case of TCE, may indirectly increase cancer risk), due to TCE exposure at Site LF-13 was modeled as

$$R = 1 - (1 - U_G R_G)(1 - R_C) , \tag{22}$$

in which U_G was just defined; R_G and R_C were defined by Equations 16a and 20a, respectively; and correlations between R_G and R_C were incorporated (see Section 2.6).

2.5. Risk Characterization

Increased health risk and related JUV associated with residential exposure to TCE from ground water at Site LF-13 on Beale Air Force Base in California was characterized quantitatively using notation similar to those used in the Phase-1 report (Daniels et al., 2000). Specifically, increased individual risk R defined by Equation 22 was evaluated using established methods (Bogen, 1995; Bogen and Spear, 1987; NRC, 1994) to obtain mean and upper-bound values of the conditional expectations $\overline{R}$ and $\langle R \rangle$, where the cdf for $\overline{R}$ represents uncertainty in risk to a (hypothetical) person at a population-average level of risk relative to others, and the cdf of $\langle R \rangle$ represents interindividual variability in the expected values of risk predicted for different people. A subscript p ($0 \le p \le 1$) on either of these conditional expectations is used to denote a $100p$th percentile at which the corresponding cdf is evaluated, while $R_{u,v}$ ($0 \le u \le 1$, $0 \le v \le 1$) is used to denote joint $100u$th-uncertainty and $100v$th-variability percentiles with respect to JUV in R. Estimates of $R_{u,v}$ were obtained jointly conditional on one of three upper bounds u (0.50, 0.95 or 0.99) with respect to aggregate uncertainty, and on one upper bound ($v = 0.99$) with respect to aggregate variability. These $R_{u,v}$ estimators

characterize median and upper uncertainty bounds on risk to a person who is relatively highly at risk compared to others at risk.

The JUV-explicit estimators of individual risk obtained (involving $\overline{R}$, $\langle R \rangle$, and $R_{u,v}$) were compared to traditional point-estimates of risk $\hat{R}_{RME}$ and $\hat{R}_{High}$ taken from Daniels et al. (2000). The $\hat{R}_{RME}$ estimate was calculated entirely analytically, using regulatory default values for all input variates where available; where default values were not available, expected values were used for all uncertain variates, and upper/unlikely bounds (e.g., 95th percentile values) were used for all heterogeneous variates. The $\hat{R}_{High}$ estimate was similarly calculated using only upper/unlikely bounds for all input variates.

Also of potential interest to stakeholders and decision makers are corresponding estimates of population risk, that is, of the uncertain total number N of additional cases of TCE-induced cancer or noncancer associated with population exposure to risk R. For an exposed population of total size n, N has an expected value of $\langle N \rangle = n \langle \overline{R} \rangle$, and the probability p_0 that there will be zero additional cases (and consequently zero health benefit from efforts to reduce R) is well approximated by the integral of the conditional Poisson likelihood function

$$p_0 \quad \approx \quad \int_0^1 e^{-n\overline{R}} f_{\overline{R}}(r)\, \mathrm{d}r \quad , \tag{23}$$

in which $f_{\overline{R}}(r)$ is probability density function of the uncertain conditional expectation $\overline{R}$ referred to above (in reference to R defined by Equation 22), and the compound-Poisson rate $(n\overline{R})$ incorporates this same conditional expectation (see Bogen and Spear, 1987; NRC, 1994; Bogen, 1995).

2.6. Data Analysis and Computation

Uncertain cancer potency U_{pot} for each animal-bioassay data set was calculated using a computationally efficient, non-asymptotic, analytic-bootstrap method previously described (Bogen, 1994). Briefly, potency was estimated for each bioassay data set using least-squares polynomial-regression fits of a polynomial in LTWA effective bioassay dose D_G to 500 simulated values of $-\ln[1-P(D_G)]$, based on observed

tumor-occurrence rates $P(D_G)$ at each level of D_G used, and under the constraint that all fitted polynomial coefficients are nonnegative. The polynomial degree was specified in the usual way, as previously described (Anderson et al., 1983). Uncertainty in U_{pot} reflected by each data set was then modeled as the empirical distribution corresponding to the 500 resulting fitted values of potency, defined as the linear coefficient in dose.

Estimated parameter and asymptotic SD values were obtained for a lognormal model (Equation 17) fit to mouse TBARS-vs.-A_C data by Levenberg-Marquardt minimization of X^2, the sum of weighted squared deviations of observed from predicted values; corresponding goodness-of-fit was assessed as $Prob(X^2 > \chi^2)$ for χ^2 distributed as chi-square with degrees of freedom (df) equal to the number of data points minus the number of estimated parameters (Press et al., 1992). The weight used for each A_C level was the corresponding value s^{-2}, where s = the SD of raw TBARS measures calculated from the SD of the corresponding mean TBARS value reported by (Larson and Bull, 1992a).

Equation 18 was evaluated numerically in order to calculate the risk of cytotoxic response, $F_C(D_{Ca})$, as a function of D_{Ca} over the interval $0 \leq F_C(D_{Ca}) \leq 1$. This numerical evaluation made use of the fact that the exponentiated expression in Equation 18 involving U_{t1} and U_{t2} represents a linear function of correlated t-distributed variates, namely: $w_0 + w_1 U_{t1} + w_2 U_{t2}$, where $w_0 = \hat{\mu} + \hat{\sigma}\{\Phi^{-1}[(Y_{sig}-Y_0)/100]\}$, $w_1 = s_{\hat{\mu}}$, $w_2 = s_{\hat{\sigma}}\{\Phi^{-1}[(Y_{sig}-Y_0)/100]\}$, and a correlation r is assumed between U_{t1} and U_{t2}. Exact and corresponding approximate expressions are available for the distribution of linear functions of independent t-distributed variates (Ruben, 1960; Patil, 1965; Ghosh, 1975; Walker and Saw, 1978; Chaubey and Mudholkar, 1982; Ojo, 1988; Singh, 1990), but not for correlated t-distributed variates. The distribution of any linear function of correlated t-distributed variates is derived in Appendix 2. In particular, the weighted sum of two t-distributed variates U_{t1} and U_{t2} with correlation r and with degrees of freedom f_1 and f_2, respectively, is given by:

$$F(v) = \int_0^1 Prob(T_{f_1+f_2} \leq v\phi(x)) [d\beta(x,{f_1}/{2},{f_2}/{2})/dx]\, dx \quad , \quad \text{where} \tag{24a}$$

$$\phi^2(x) = \frac{(f_1+f_2)x(1-x)}{f_1(1-x)w_1^2 + f_2 x w_2^2 + 2r w_1 w_2 \sqrt{f_1 f_2 x(1-x)}} \quad , \tag{24b}$$

and where in Equation 24a, $T_{f_1+f_2}$ denotes a Student's t-distributed variate with f_1+f_2 degrees of freedom, and β denotes the incomplete beta-function ratio corresponding to the specified arguments (see Abbreviations and Notation). To approximate $F_C(D_{Ca})$ over relevant risk range ($\sim 10^{-10}$ to $\sim 10^{-2}$), the linear-quadratic function $Log_{10}(risk) = a+bx+cx^2$ was fit by unweighted least squares to points on $F_C(D_{Ca})$ calculated within this risk range, where $x = Log_{10}(D_{Ca})$. The following second-order approximation was then obtained using Equation 21 (in Section 2.4) to estimate corresponding conditional risks under MA_C (see Appendix 3.H, pp. H-9 to H-10):

$$ER_C \approx E_1R_C + \frac{10^a}{2}\Pi^{b+cL}\big(\lambda + b(b-1) + 2cL[2(b+cL)-1]\big)\big(\gamma_{X_1}^2 + \gamma_{X_2}^2\big) \ . \tag{25}$$

Equations 20b and 21 imply, in particular, that $\overline{R_C}$ is estimated by Equation 25 in which: E denotes expectation with respect to variability only; a, b, and c are the linear-quadratic coefficients defined above; $\Pi = KUV_1V_2$; $K = \langle\overline{D_C}\rangle$; $U = U_{tdyn}\big(1+U_{chron}\big)\overline{B_C}$; $V_1 = V_{tdyn}$; $V_2 = \langle B_C\rangle$; $\lambda = 2c/\ln 10$ (where ln denotes natural logarithm); $L = \log_{10}\Pi$; $\gamma_X = (\sigma_X/EX) =$ the coefficient of variation of the subscripted variate X; $X_1 = V_1$; and $X_2 = V_2$. Likewise, Equations 20c and 21 imply that $\langle R_C\rangle$ is estimated by Equation 25 in which: E denotes expectation with respect to uncertainty only; $\Pi = [KVU_1(\sum_i U_i)]$ for $i = \{ing, inh, der\}$; $V = V_{tdyn}\langle B_C\rangle$; $U_1 = U_{tdyn}\big(1+U_{chron}\big)$; $U_i = f_{C,i}\overline{B_C}$; $L = \log_{10}\Pi$; K, λ and γ were defined above; $X_1 = U_1$; and $X_2 = \sum_i U_i$ for $i = \{ing, inh, der\}$, noting that $EX_1 = EX_2 = 1$ and that $\sigma_{X_2}^2 = \Big(\sum_i f_{C,i}^2\Big)\sigma_{\overline{B_C}}^2$. Equation 24 was evaluated taking all moments with respect to the same distributed characteristic (i.e., uncertainty or variability, corresponding to E as defined above), using estimates for a, b, and c, and for indicated variate moments, that are implied by assumptions used (see Appendix 1) or are reported in Results (Section 3).

Monte-Carlo methods were used to generate sample values for each of (say, k) distributed variates involved in a given calculation. Specifically, systematic Latin-hypercube sampling was used to simulate n_{sam} samples of each required set of k variates, where k was determined by the equation(s) being evaluated, and a method

(Iman and Conover, 1982) was used to obtain rank-correlated sample vectors, each with a rank-correlation matrix M_i ($i = 1,...,k$) not significantly different ($p_{adj} > 0.05$) from a specified target matrix T, which by default was a $k \times k$ identity matrix modified to reflect correlations specified below. A value of $n_{sam} = 2000$ was used unless otherwise specified. The k differences between M_i and T were each assessed using an asymptotic chi-square test (Jennrich, 1970), and the p-value from each test was adjusted (to p_{adj}, to account for k independent tests) using Hommel's Bonferroni-type procedure (Wright, 1992). Typically, $\text{Min}(p_{adj}) > 0.95$; occasional sample vectors not satisfying $p_{adj} > 0.01$ were rejected. Each simulation was repeated n_{sim} times, a grand AM and its CV (denoted CVM, where $CVM = CV[n_{sim}]^{-0.5}$) from the n_{sim} cdf-specific AMs, and the AM and CVM were calculated for each ith set of n_{sim} cdf-specific order statistics (i.e., cdf-abscissa values), where $i = 1,...,n_{sam}$. The calculated CVM values reflect simulation quality by indicating the relative size of Monte-Carlo sampling error produced for estimators of interest conditional on the values of n_{sam} and n_{sim} used.

Target rank-correlation values or matrices were estimated for all sets of correlated variates noted or implied above, namely, the sets: $\{V_{fm,inh}, V_{alv}\}$, $\{V_{fd}, V_W\}$, $\left\{ \langle B_{G,ing} \rangle, \langle B_{C,ing} \rangle, \langle B_{G,inh} \rangle, \langle B_{C,inh} \rangle, \langle B_{G,der} \rangle, \langle B_{G,der} \rangle \right\}$, $\{ \overline{B_G}, \overline{B_C} \}$, and $\{U_{t1}, U_{t2}\}$ (see Results, Sections 3.1 and 3.2). These correlations were used, respectively, to evaluate: $D_{G,inh}$ and $D_{C,inh}$ in Equations 6 and 15b; $D_{C,P}$ in Equations 15a-b; $\langle B_{MA,P} \rangle$ in Equations 16c and 20c; $\overline{B_{MA}}$ in Equations 16b and 20b, and $F_C(D_{Ca})$ in Equation 18. In calculations to estimate correlations involving $\langle B_{MA,P} \rangle$, values of $n_{sim} = 500$ and $n_{sam} = 50$ were used.

Correlations involving all the variate sets listed above were used in nested (i.e., two-dimensional) Monte-Carlo evaluations of Equation 22 (which, in this case, refers to Equations 16a and 20a) that were performed to estimate $R_{u,v}$. For these nested calculations, values of $n_{sim} = 100$ and $n_{sam} = 999$ were used, with n_{sam} used to simulate all uncertain variates, and then used again to simulate all heterogeneous variates conditional on each of the n_{sam} simulated sets of uncertain variates. Because $\langle B_{MA} \rangle \approx \langle B_{MA,P} \rangle$ and $\overline{B_{MA}} \approx \overline{B_{MA,P}}$ for all pathways P (see Results, Section 3.1), $\langle B_{MA} \rangle$ and $\overline{B_{MA}}$ were used to evaluate Equations 16b and 20d, rather than $\langle B_{MA,P} \rangle$ and $\overline{B_{MA,P}}$. However, to evaluate Equations 16a-c, 20a-c, 22, and 25, pathway-specific dose

correlations noted above were applied (as applicable) to EB_{MA} and B_{MA} to regenerate the pathway-specific variates involved in these equations.

All calculations were performed on a PowerMac G4 computer using the programs *Mathematica*®4.0 (Wolfram, 1999) and *RiskQ* 4.0 (Bogen, 2000). Documentation of these calculations appears in Appendices 3.A through 3.I, in which calculations and related comments are organized by topic. Appendices 3.A (Concentration), 3.B (Intakes), and 3.C (Fraction of Lifetime at One Local Residence) all document the derivation or re-derivation of exposure-related input variates explained in Daniels et al. (2000), which were used to calculate TCE exposures as explained above (Section 3.2). Appendices 3.D (Effective Genotoxic Dose) and 3.E (Effective Cytotoxic Dose) document the calculation of corresponding biologically effective (TCE or TCA) doses. Note that calculations pertaining to the definition or characterization of variates V_W, V_{Vmax}, $V_{fm,ing}$, V_{fd}, $(f_{deq}/V_{t,P})$ and V_e all appear in Appendix 3.E. Appendix 3.F (Effective Dose Correlations) documents calculations made to estimate rank correlations among MA- and pathway-specific normalized biologically effective doses. Appendix 3.G (Potency) documents all calculations made pertaining to modeled dose-response under both mechanisms of carcinogenic action considered (MA_G and MA_C). Appendix 3.H (TCE Risk) documents all calculations made pertaining to corresponding predicted risk. Note that calculations pertaining to the definition of variates U_{chron}, U_{tdyn}, and V_{tdyn} appear in Appendix 3.H. Finally, Appendix 3.I (Functions Used) briefly describes all *Mathematica*® and *RiskQ* functions used to carry out calculations documented in Appendices 3.A-3.H. More detailed explanation of *Mathematica*®, *RiskQ*, and JUV analysis is beyond the scope of this report, and is provided in references cited.

All constants and variates defined in this report that were used as input to estimate risk, as described above in Sections 2.1 - 2.6, are summarized in the following table (Table 2).

Table 2. Constants and variates used as input for unified TCE risk assessment.[a]

Input type[b]	ID[b]	Description	Unit	1st used in or near Equation(s) (No.)	Distribution type, parameter value(s)[c]			Reference(s)[d]
					Dist	μ	σ	
K	fTCA	Fraction of total TCE intake metabolized to TCA	unitless	10		0.33		Allen & Fisher (1993)
K	MW	TCA to TCE molecular-weight ratio	unitless	10		1.228		Bogen & Gold (1997)
P	Y_0	Background relative TBARS level in mouse liver	unitless	17, 18		40		Larson & Bull (1992)
P	Y_{sig}	Significantly elevated relative TBARS level	unitless	18		49.4		Larson & Bull (1992)*
P	$\hat{\mu}$, $s_{\hat{\mu}}$	Estimated location parameter of lognormal cytotoxicity dose-response model, and its SD	unitless	18		3.05	0.0920	Calculated from data of Larson & Bull (1992)
P	$\hat{\sigma}$, $s_{\hat{\sigma}}$	Estimated shape parameter of lognormal cytotoxicity dose-response model, and its SD	unitless	18		0.732	0.176	Calculated from data of Larson & Bull (1992)
P	$\hat{a}$	Estimated log-linear-regression intercept parameter used to extrapolate cytotoxicity dose-response	unitless	23		-7.60		Calculated (see text)
P	$\hat{b}$	Estimated log-linear-regression slope parameter used to extrapolate cytotoxicity dose-response	unitless	23		3.68		Calculated (see text)
U	pot	Carcinogenic potency of TCE assuming a genotoxic mechanism of action	kg d mg^{-1}	16	Emp	3.7×10^{-4}		Calculated from data on 7 studies (see text)
U	tdyn	Uncertainty factor for interspecies extrapolation of toxicodynamically equivalent effective dose	unitless	16, 20	LN	0	1.6	EPA (1998)* Slob & Pieters (1999)*
U	chron	Uncertainty factor for extrapolation of acute to chronic cytotoxic dose	unitless	19	LN	ln 5	ln 2.12	NCI (1976)* Slob & Pieters (1999)*
U	t1	Normalized estimation error in $\hat{\mu}$	unitless	18	T[f]	14		Assumed
U	t2	Normalized estimation error in $\hat{\sigma}$	unitless	18	T[f]	14		Assumed
V	W	U.S. adult male and female body weight	kg	2	LN[e]	4.24	0.221	Finley (1994) CalEPA (1996, p. 10-7)

43

Table 2. Constants and variates used as input for unified TCE risk assessment (continued).[a]

Input type[b]	ID[b]	Description	Unit	1st used in or near Equation(s) (No.)	Distribution type, parameter value(s)[c] Dist	μ	σ	Reference(s)[d]
V	alv	Normalized variability in alveolar ventilation rate, independent of variability in V_W	unitless	2, 3	LN	-0.0409	0.286	CalEPA (1996, p. 3-31)*
V	Pb	Human blood:air partition coefficient for TCE	L_{air}/L_{blood}	7	N	10.2	1.6	Fisher (1998)*
V	liv	Human blood flow to liver = 26%×(15.0/12.9)×V_{alv}	mL h^{-1}	8	(see V_{alv})			Allen & Fisher (1993)
V	V_{max}	Normalized variability in the maximum rate of TCE metabolism, independent of variability in V_W	unitless	8	LN	-0.152	0.551	Lipscombe et al. (1998)*
V	fd	Fraction of V_W corresponding to apparent volume of distribution for TCA	L kg^{-1}	10	U[e]	0.052	0.152	Allen & Fisher (1993)*
V	e	Normalized variability in TCA elimination rate V_{ke}, independent of variability in V_W	unitless	10		-0.152	0.551	
V	tdyn	Variability factor modeling intraspecies differences in sensitivity (i.e., in toxicodynamically equiv. dose)	unitless	16, 20	LN	0.700	2.33	EPA (1998)* Slob & Pieters (1998)*

[a]Constants and variates listed are those defined in this report and used or implied in Equations 1 - 21 as inputs to risk estimation. Variates defined by Equations 1 - 21 are not repeated in this table. All other variates that were used to estimate risk are either defined in Daniels et al. (2000), or are defined in Equations 1 - 21 in terms of constants and variates listed in this table.

[b]Input types: K = constant, P = estimated (hence, constant) parameter value, U = uncertain variate, V = heterogeneous variate (i.e., values pertain to different individuals at risk). ID = the subscript that appears in the text on a K-, U- or V-type input; ID = the symbol used in the text to denote a P-type input.

[c]Dist specifies distribution type: LN = lognormal, N = standard normal, T = Student's T, U = uniform, Emp = calculated empirical, Blank = not applicable. The values $\{\mu, \sigma\}$ = {the estimated/assigned value, (if applicable) the SD} of for K- or P-type inputs; otherwise $\{\mu, \sigma\}$ denote (for the specified Dist): {ln GM, ln GSD} (LN), {AM, SD} (N), {df, –} (T), and {min, max} (U).

[d]An asterisk signifies that the value or approach cited was modified slightly or generalized for use in this report.

[e,f]These variates assumed to have a rank correlation equal to: -0.50^e, 0.294^f.

44

3. RESULTS

Resulting estimates of biologically effective dose TCE contamination at Site LF-13 are presented below in Section 3.1, followed (in Section 3.2) by estimated dose-response relations obtained. Finally, Section 3.3 provides a characterization of corresponding risks and associated JUV estimated using the systematic probabilistic framework adopted for this study, as well as a comparison of these estimates with point-estimates of risk for Site LF-13 obtained using traditional methods.

3.1 Biologically Effective Dose

The cdfs obtained to characterize variability in the limiting fraction $V_{fm,ing}$ of low-level ingested TCE that is metabolized, and in the corresponding limiting fraction $V_{fm,inh}$ (= $V_{fm,der}$) of low-level respired or dermally absorbed TCE that is metabolized, are shown in Figure 3. The variates $\{V_{fm,inh}, V_{alv}\}$ were found to have an approximate rank correlation of -0.75 (CVM = 0.33%). Although not used in calculations performed in this study, the rank correlation between variates $\{V_{fm,ing}, V_{fm,inh}\}$ was found to be ~0.83 (CVM = 0.30%).

The JUV-expectation of genotoxic effective dose, $\langle \overline{D_G} \rangle$, was found to be 5.93×10^{-5} mg kg^{-1} d^{-1} (CVM < 1%), with corresponding pathway-specific dose fractions: $f_{G,ing} = 0.843$, $f_{G,inh} = 0.039$, and $f_{G,der} = 0.118$. The JUV-expectation of cytotoxic effective dose, $\langle \overline{D_C} \rangle$, was found to be 0.0269 mg L^{-1} (CVM < 1%), with corresponding pathway-specific dose fractions: $f_{C,ing} = 0.604$, $f_{C,inh} = 0.312$, and $f_{C,der} = 0.084$.

The cdfs obtained for the three pathway-specific expectations with respect to uncertainty in normalized effective genotoxic dose ($\langle B_{G,P} \rangle$ for P ={ing, inh, der}, shown as three bold curves), and for the corresponding three pathway-specific expectations with respect to variability in normalized effective genotoxic dose ($\overline{B_{G,P}}$, three light curves), are plotted in Figure 4a. The figure shows that the three pathway-specific curves that comprise each set of (bold or light) curves are virtually indistinguishable. The cdfs obtained for the three pathway-specific expectations with respect to uncertainty in normalized effective cytotoxic dose ($\langle B_{C,P} \rangle$, three bold curves), and for the corresponding three pathway-specific expectations with respect to variability in

45
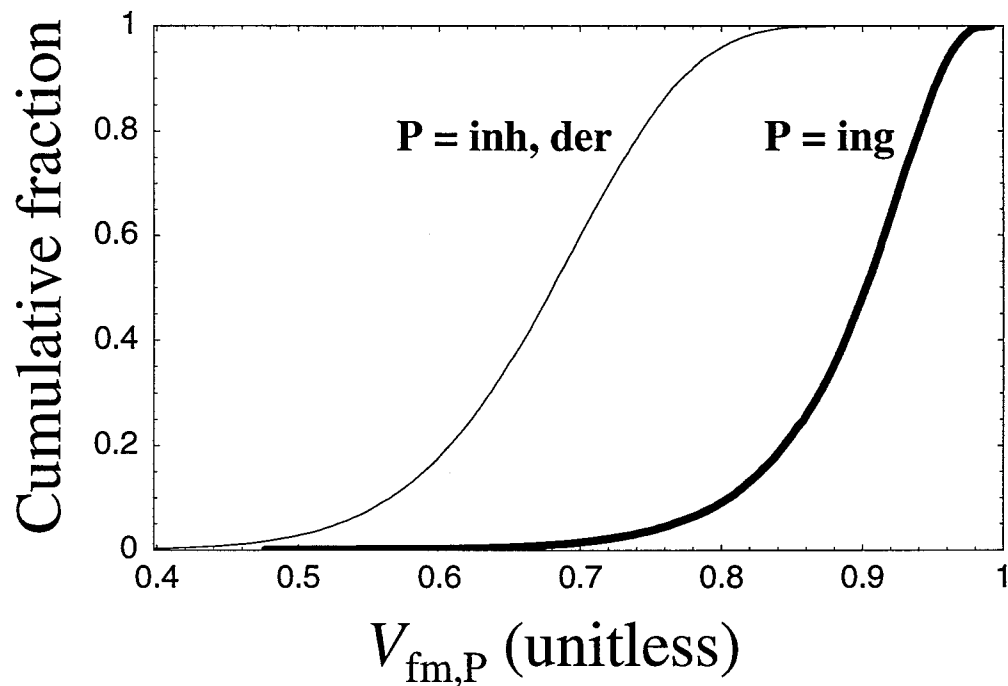
Figure 3. Cumulative distribution functions characterizing interindividual variability in limiting metabolized fractions $V_{fm,P}$ of low-level TCE absorbed via different exposure pathways P, where P = {ing, inh, and der} for {ingestion, ihalation, and dermal} pathways, respectively.
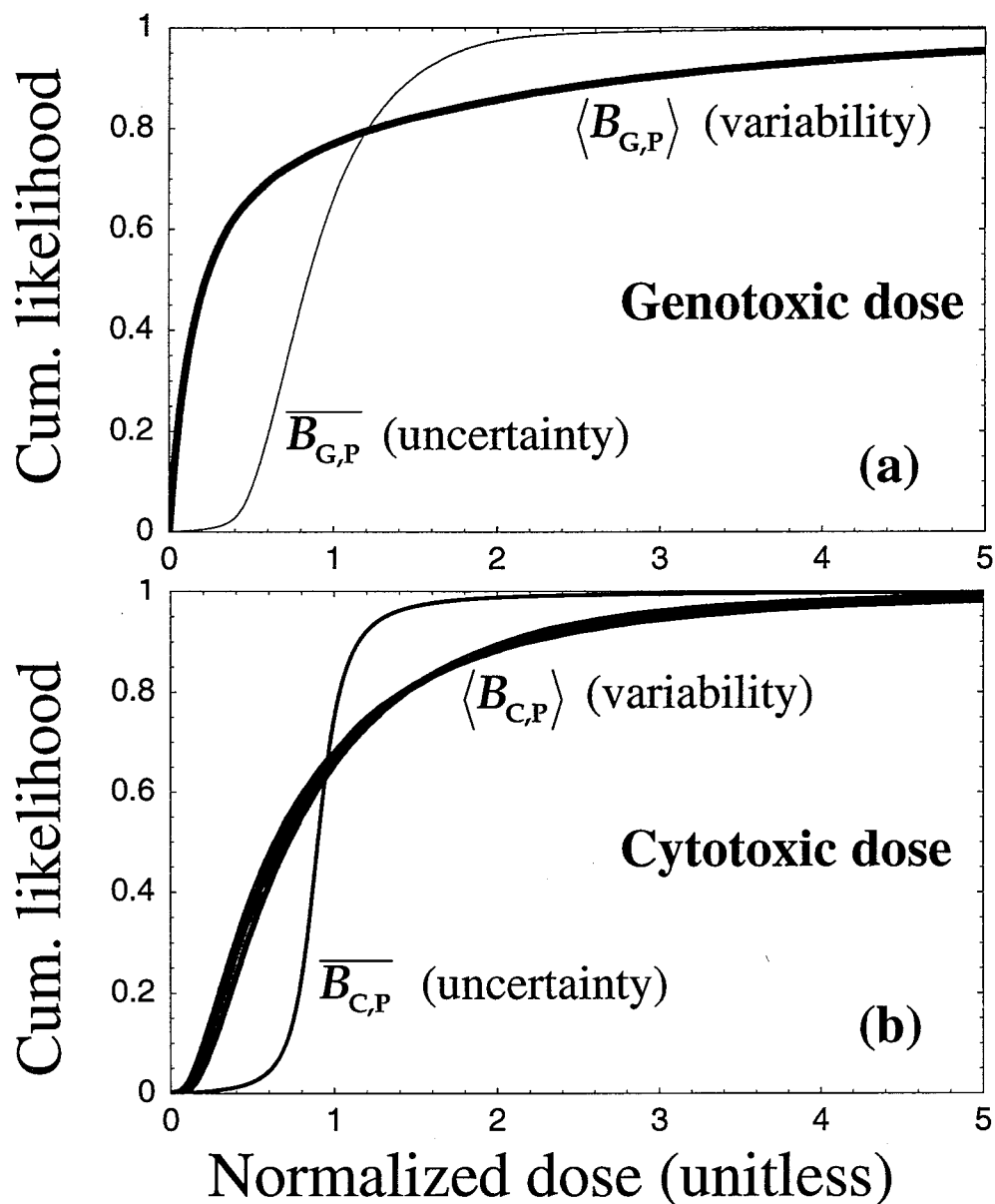
Figure 4. Expectations with respect to uncertainty vs. variability in normalized effective **(a)** genotoxic dose ($B_{G,P}$), and **(b)** cytotoxic dose ($B_{C,P}$), shown for each (P = ingestion, inhalation, and dermal) exposure pathway considered. Each corresponding cumulative distribution function (cdf) shown characterizes normalized interindividual variability in values of expected risk (bold curves), or normalized uncertainty in the population-average value of risk (light curves) predicted for hypothetical residents exposed to TCE from ground water at Site LF-13. Three bold and three light exposure-pathway-specific curves appear in each plot, but in each set these curves very nearly coincide (except for slight divergence among bold curves in plot b). All the cdfs were normalized to have an arithmetic mean value of one.

normalized effective cytotoxic dose ($\overline{B_{C,P}}$, three light curves), are plotted in Figure 4b. The figure shows that the three pathway-specific curves that comprise $\overline{B_{C,P}}$ (light curves) are virtually indistinguishable, while those comprising $\langle B_{C,P} \rangle$ (bold curves) are nearly so. Thus $\langle B_{MA} \rangle \approx \langle B_{MA,P} \rangle$ and $\overline{B_{MA}} \approx \overline{B_{MA,P}}$ for all pathways P, which justifies the exclusive reliance on $\langle B_{MA} \rangle$ and $\overline{B_{MA}}$ for calculations described in Methods. The variates $\{\overline{B_G}, \overline{B_C}\}$ were found to have an approximate rank correlation of 0.49 (CVM = 0.67%), and rank-correlation and corresponding CVM matrices obtained for the six $\langle B_{MA,P} \rangle$ variates are listed below in Table 3.

3.2 Dose-Response

The cdfs characterizing estimation error (uncertainty) in cancer potency U_{pot} estimated for each of seven animal-bioassay data sets considered are shown in Figure 5a; the corresponding weighted-average cdf based on weights indicated in Table 1 is shown in Figure 5b.

The fit of the lognormal model specified by Equation 17 to mouse TBARS-vs.-A_C data is shown in Figure 6. The model fit the experimental data reasonably well ($X^2 = 13.6$, df = 14, p = 0.48). The two corresponding parameter estimates (±1 SD) $\{\hat{\mu} = 3.05 \pm 0.0920, \ \hat{\sigma} = 0.732 \pm 0.176\}$ and their estimated correlation coefficient ($r = 0.294$) were obtained from this fit. The corresponding risk function $F_C(D_{Ca})$ calculated using Equations 24a-b is shown in Figure 7a. Figure 7b shows the risk function $F_C(D_{Ca})$ replotted on a log-log scale, together with an approximating linear-quadratic equation fit to $F_C(D_{Ca})$ as described in Section 2.6: $Log_{10}(\text{risk}) = a + bx + cx^2$, where $x = Log_{10}(D_{Ca})$, $a = -7.238$, $b = 2.466x$, and $c = 0.4699x^2$. The comparison shown in Figure 7b demonstrates that the linear-quadratic equation provides an excellent approximation to $F_C(D_{Ca})$ over a substantial risk range ($\sim 10^{-10}$ to $\sim 10^{-2}$) relevant to the present analysis. Figure 7b also shows how the slope of the $F_C(D_{Ca})$ function plotted on a log-log scale changes abruptly at a risk level of $\sim 10^{-6}$, implying that a Monte Carlo approach would not be practical as an alternative to using Equations 24a-b in order to evaluate levels of cytotoxic risk $F_C(D_{Ca}) \leq \sim 10^{-6}$, based on Equation 18.

Table 3. Rank correlations among uncertainty-expectations of normalized biologically effective doses.[a]

$\langle B_{MA,P} \rangle$ variate		$\langle B_{MA,P} \rangle$ variate										
		MA	P	MA	P	MA	P	MA	P	MA	P	MA	P
variate		G	ing	C	ing	G	inh	C	inh	G	der	C	der
MA P	G ing	1		0.23		0.88		0		0.89		0	
MA P	C ing	0.23		1		0		0.42		0		0.51	
MA P	G inh	0.88		0		1		0.19		0.92		0.035	
MA P	C inh	0		0.42		0.19		1		0.077		0.65	
MA P	G der	0.89		0		0.92		0.077		1		0.18	
MA P	C der	0		0.51		0.035		0.65		0.18		1	

[a]Estimated values of the Spearman rank correlation coefficient (r, shown with two significant digits) based on Monte-Carlo evaluation of the uncertainty-expectation of Equation 5c based on Equations 6 and 15a-b, where $n_{sam} = 500$ and $n_{sim} = 50$. For all r-values listed, SDM < 0.0025 where SDM = $(n_{sim})^{-1/2}SD(r)$ and SD(r) denotes the SD of the n_{sim} estimates of r obtained.

3.3 Predicted Risk

The individual risks predicted in this study correspond to the assumption that the assumed TCE concentration in ground water beneath Site LF-13 (~22 ppb, as of 1997) remains unchanged (see Daniels et al., 2000). The cdfs obtained that characterize uncertainty in the predicted population-average value of individual risk, $\overline{R}$, and interindividual variability in expected values (i.e., "best" estimates) of individual risk, $\langle R \rangle$, are shown in Figures 8a and 8b, respectively, plotted together with corresponding CVM values. In Figures 9a-c, the cdfs $\overline{R}$ and $\langle R \rangle$ are contrasted over different ranges
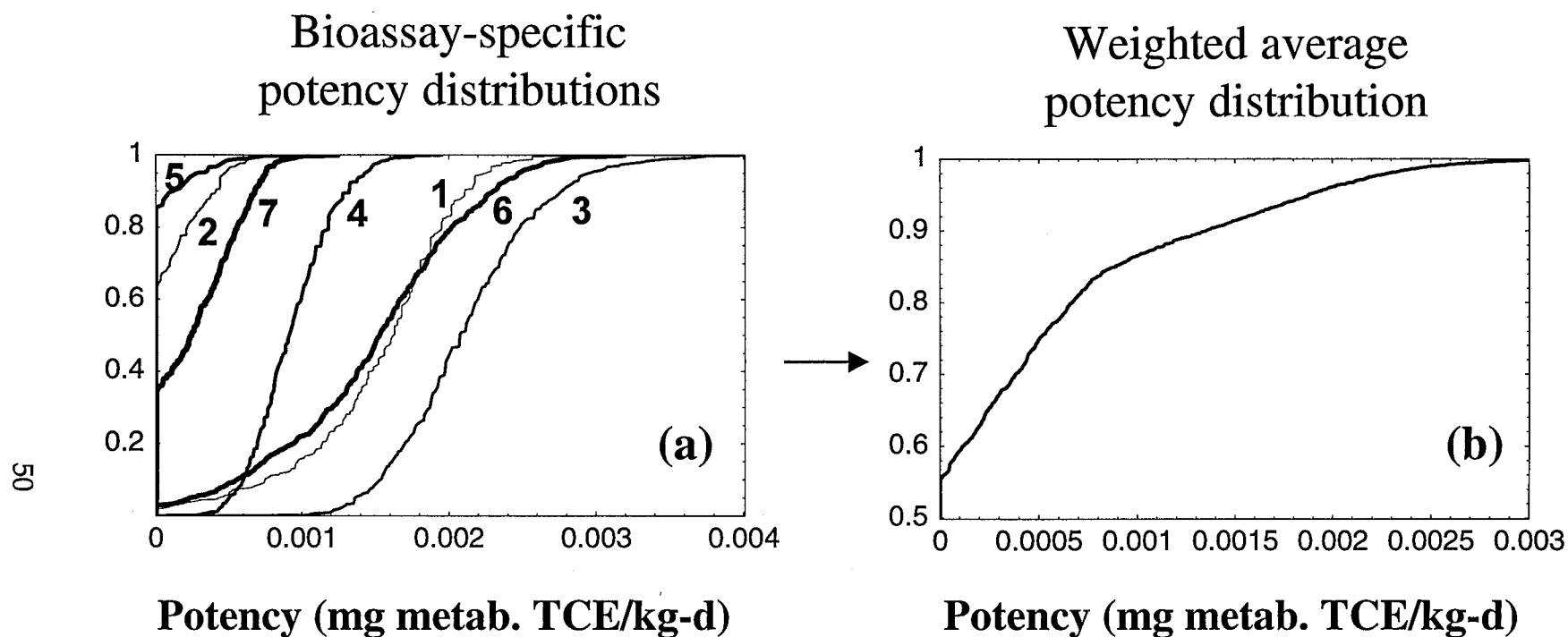
Figure 5. Estimation error (uncertainty) in cancer potency based on rodent-bioassay data. Cumulative distribution functions (cdfs) shown characterize uncertainty in potency estimates based on **(a)** individual data sets, and **(b)** a corresponding weighted average of the species/strain/sex-specific cdfs. The numbers labeling individual cdfs in (a) correspond to the study numbers listed in the first column of Table 1. The vertical axis represents cumulative probability.
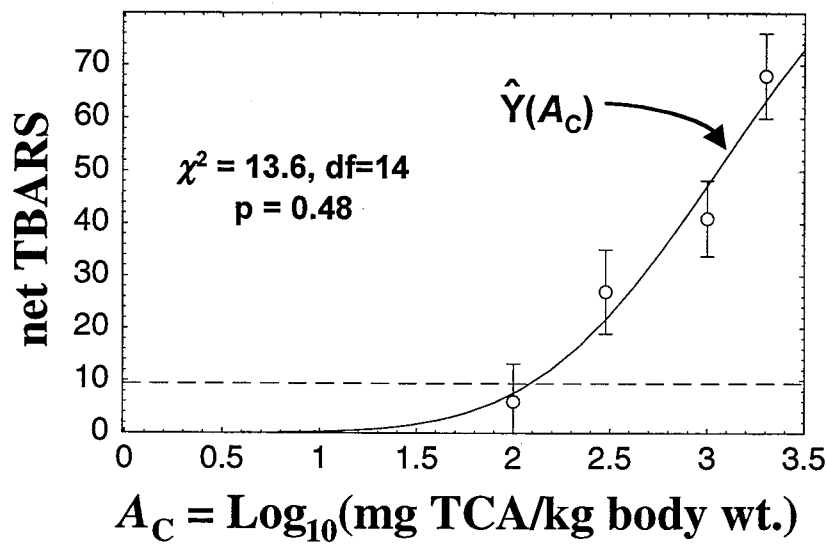
Figure 6. Fit of lognormal model to mouse cytotoxicity data from Larson and Bull (1992a). Experimental mean (±1 SD) data (open points) are shown relating administered TCA dose (A_C) to a hepatocellular lipoperoxidative index (TBARS) associated with liver cytotoxicity, from which the mean (Y_0) of measured control TBARS levels has been subtracted. The dashed line indicates the 2-tailed 95% upper confidence limit (UCL, Y_{sig}) on Y_0, minus Y_0. The lognormal model shown (curve; see Equation 17) was fit to the data.
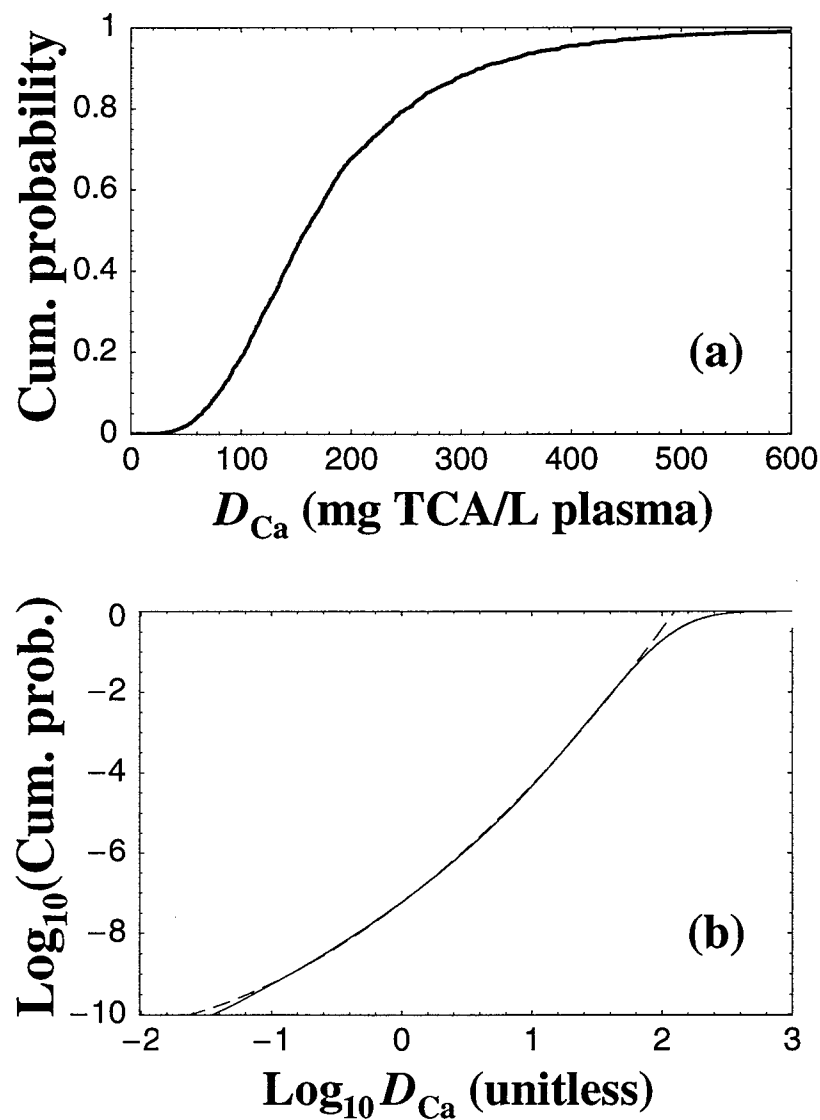
Figure 7. (a) Risk of cytotoxic response, $F_C(D_{Ca})$, as a function of acute effective administered TCA dose D_{Ca}, estimated from mouse cytotoxicity data of Larson and Bull (1992a) on TCA-induced TBARS elevation (see Figure 6). **(b)** The function $F_C(D_{Ca})$ (solid curve) replotted on a log-log scale, together with the approximating equation, $Log_{10}(risk) = -7.238 + 2.466x + 0.4699x^2$ (dashed curve), where $x = Log_{10}(D_{Ca})$.
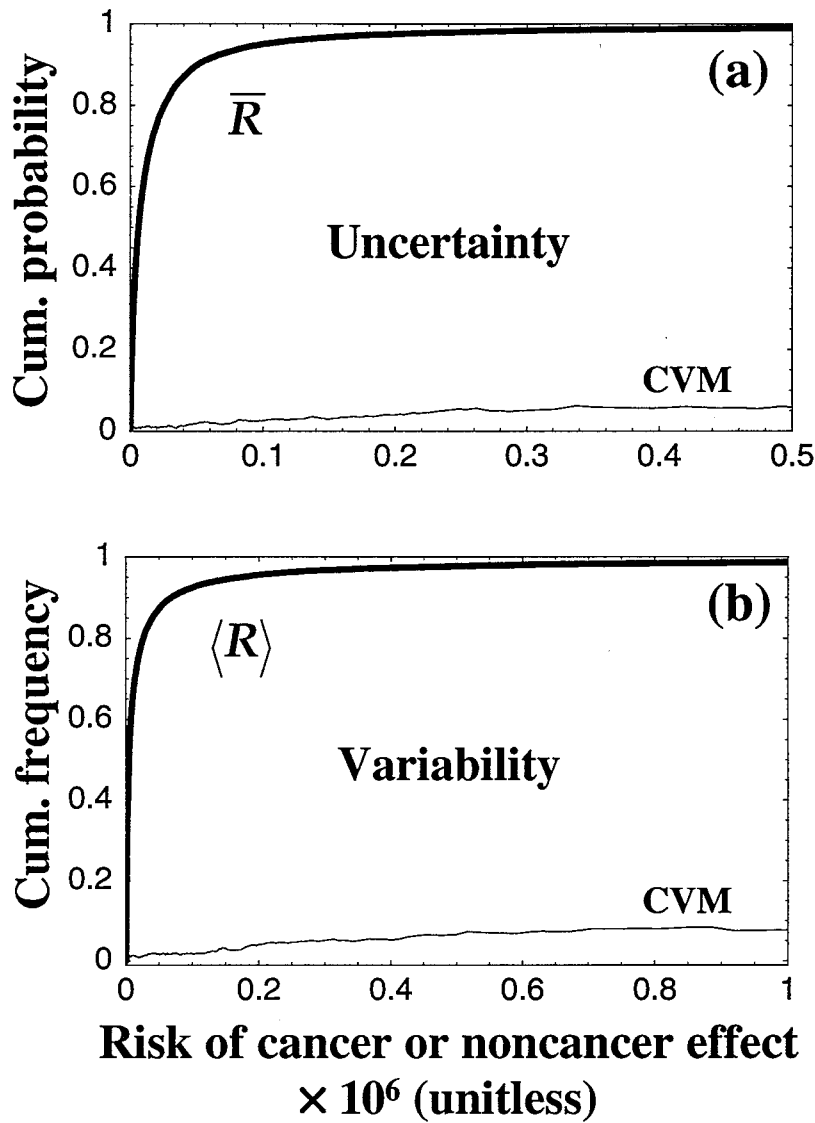
Figure 8. Uncertainty in (a) population-average risk, $\overline{R}$, and (b) interindividual variability in expected risk, $\langle R \rangle$, predicted for hypothetical individuals exposed to TCE in ground water from Site LF-13. In each plot, the Monte-Carlo relative-sampling error of the x-axis value of each point on the bold cdf curve is indicated by the corresponding y-axis value (labeled CVM) of the light curve shown. For example, from plot (a) the 99th percentile value of $\overline{R}$ is estimated to be ~0.49×10[-6], which estimate has a CVM of ~0.060, indicating a sampling error of about ±6%.
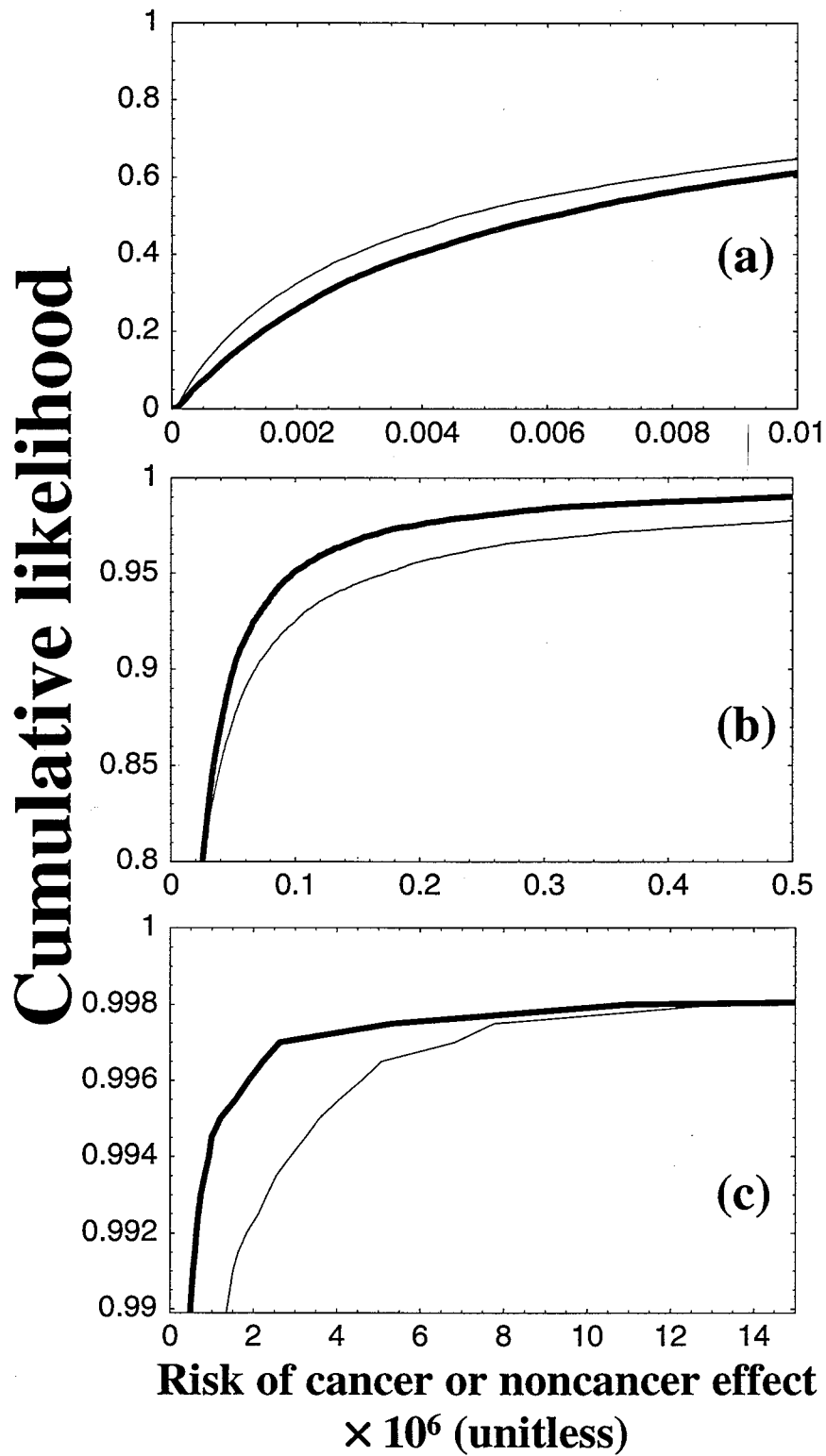
Figure 9. Comparison of $\overline{R}$ (bold curves) vs. $\langle R \rangle$ (light curves) over different ranges of predicted risk: **(a)** $R \leq 0.01 \times 10^{-6}$, **(b)** $R \leq 0.5 \times 10^{-6}$, and **(c)** $R \leq 15 \times 10^{-6}$. The relationship is not consistent over these risk ranges, and both cdfs are highly skewed.

of risk. Figures 10a-b compare cdfs with respect to variability in JUV estimators $R_{u,v}$ conditional on specified confidence bounds (u) on uncertainty, for u = 0.50, 0.95, and 0.99. The nested Monte-Carlo calculations required to estimate these three $R_{u,v}$ cdfs took a total of 10.6 h to perform.

Table 4 provides a comparison of an upper variability bound (v = 0.99) on these cdfs to mean and upper-bound values of $\overline{R}$ and $\langle R \rangle$, as well as to traditional point-estimates of risk ($\hat{R}_{RME}$ and $\hat{R}_{High}$) taken from Daniels et al. (2000). CVM values were all about 10% or less, except for a CVM value of 23% for the expected value of $\overline{R}$. These CVM values indicate reliability in the result obtained that the estimated mean and upper-bound values of $\overline{R}$ and $\langle R \rangle$ are all $\leq \sim 10^{-6}$, and that both JUV estimators are $<$ 5×10^{-5}.

Corresponding estimates of population risk (i.e., the uncertain number N of cases of cancer or TCE-induced toxicity) depend on the assumed size n of the total exposed population (including all immigrants to and emigrants from areas hypothetically served by Site LF-13 water), for an arbitrarily assumed total period equal to one average lifetime (taken to be 70 y), during which total or partial lifetime exposures would hypothetically occur (see Daniels et al., 2000). Expected population risk $\langle N \rangle$ conditional on various assumptions concerning population size n are listed in Table 5, together with estimates of the corresponding likelihood ($1-P_0$) of one or more cases, and the likelihood (P_0) of zero cases, being attributable to TCE in groundwater at Site LF-13 over the 70-y period of consideration.
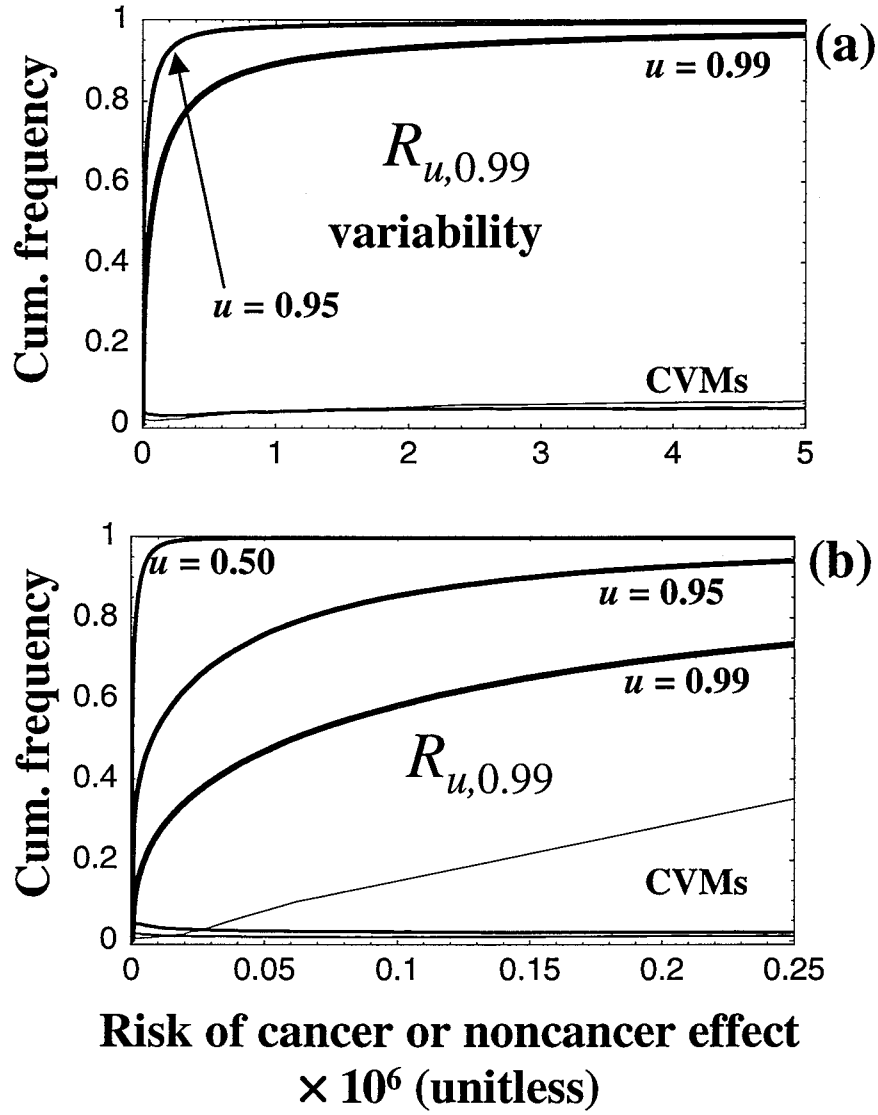
Figure 10. Comparison of estimators, $R_{u,v}$, o f joint uncertainty and variability in risk, over different ranges fo risk, where u and v refer to fractiles with respect to uncertainty and variability, respectively, and where the upper-bound value $v = 0.99$ was used. In (a), $R_{u,0.99} \leq 5 \times 10^{-6}$, an d in (b) $R_{u,0.99} \leq 0.25 \times 10^{-6}$. CVM cu rves denote corresponding relative error, as in Figure 8. The CVM that exceeds 0.20 where $R_{u,0.99} = 0.25 \times 10^{-6}$ corresponds to value of $u = 0.50$ (median uncertainty), conditional on w hich, for example, $R_{0.5,0.99} = 0.017 \times 10^{-6}$. The CVM > 0.20 therefore pertains to a very unlikely level of risk.

56

Table 4. Summary of estimated risk posed by TCE in ground water at Site LF-13.

Method Type of risk estimator	Symbol	Value ($\times 10^6$)	Error (%)*
Traditional			
Risk to "reasonably maximum exposed" indiv.	R_{RME}	61.	NA
Upper "conservative" bound (using 95%UCLs)	R_{Hi}	240.	NA
Explicit with respect to:[a]			
U in population-average risk $\overline{R}$ $n_{sim} = 10$			
Expectation (with respect to U)	$\langle \overline{R} \rangle$	0.92	69.
Upper 95% U-bound	$\overline{R}_{0.95}$	0.098	2.5
Upper 99% U-bound	$\overline{R}_{0.99}$	0.49	6.0
V in expected risk $\langle R \rangle$ $n_{sim} = 10$			
Population-average (with respect to V)	$\overline{\langle R \rangle}$	0.23	43.
Upper 95% V-bound	$\langle R \rangle_{0.95}$	0.17	2.8
Upper 99% V-bound	$\langle R \rangle_{0.99}$	1.4	8.1
Joint U and V in risk: R $n_{sim} = 100$			
Upper (95%-99%) (U,V)-bound	$R_{0.95,0.99}$	2.0	4.2
Upper (99%-99%) (U,V)-bound	$R_{0.99,0.99}$	37.	5.3

[a]NA = not applicable, U = uncertainty, V = interindividual variability. Error = coefficient of relative variation of the mean of n_{sim} U-, V-, or JUV-related estimates, each based on n_{sam} = 2000 Monte-Carlo-simulated sets of model-parameter values, except JUV estimators, which used n_{sam} = 999.

Table 5. Population risk associated with multipathway exposures to TCE-contaminated ground water at Beale Air Force Base in California.[a]

Total exposed population over 70 y n	Exposed population during 7.6 y n_{res}	Prob($N = 0$)[b] P_0	Prob($N \geq 0$)[b] $1 - P_0$	Expected value[b] of N $\langle N \rangle = n \times \langle \overline{R} \rangle$
100	11	0.999913	0.000089	0.000092
1,000	109	0.999423	0.00058	0.00092
2,000	217	0.999166	0.00083	0.0018
10,000	1,086	0.998468	0.0015	0.0092
30,000	3,257	0.997568	0.0024	0.027
100,000	10,857	0.995219	0.0048	0.092
1,000,000	108,571	0.974127	0.026	0.92
99,796,500	10,835,049	0.500000	0.50	91.

[a]N = population risk, i.e., the predicted number of cases (i.e., individuals with) a cancer or noncancer endpoint due to exposure to TCE from Site LF-13; n = the total number of individuals assumed to incur the population-average risk $\overline{R}$ within a 70-y period of consideration; n_{res} = mean number of exposed people at any given moment assumed to be served by ground water from Site LF-13, assuming a mean 7.6-y duration of residence (see Daniels et al., 2000).

[b]Probabilities P_0 are shown rounded to 6 decimal places; complimentary probabilities ($1-P_0$) are shown rounded to 2 decimal places; and exposed population n_{res} is shown rounded to the nearest integer; no more than 2 significant digits are implied in estimates listed in columns 2 through 5.

4. DISCUSSION

A systematic probabilistic framework was used to estimate the aggregate risk of cancer and noncancer endpoints for hypothetical future residents exposed to TCE from ground water at an inactive landfill site at Beale U.S. Air Force Base. The framework used here differs from previous approaches (Baird et al., 1996; Butterworth and Bogdanffy, 1999; Carlson-Lynch et al., 1999; Gaylor et al., 1999; Lewis, 1993; Slob and Pieters, 1998) in that it is the first to provide an integrated, consistent treatment of: (i) cancer as well as noncancer endpoints, (ii) two disparate yet plausible mechanisms of carcinogenic action in the case TCE (genotoxic vs. cytotoxic), (iii) pharmacokinetic considerations, and (iv) quantitative analysis of JUV in model inputs and corresponding characterized risk. The framework incorporates some of the probabilistic methods suggested previously, but modifies others in important and/or fundamental ways. In particular, the human-variability factor V_{tdyn} (that has an AM of 1) and the acute-to-chronic uncertainty factor $(1+U_{chron})$ (that by definition is >1) used here differs fundamentally from analogous factors recommended by Slob and Pieters (1998), in order for these factors to fit logically within a systematic probabilistic framework that addresses points (i)-(iv) above. The proposed framework is also the first ever to address both issues (i) and (ii) for the purpose of quantitative risk assessment, independent of issues (iii) and (iv). Specifically, the systematic probabilistic framework used in this case study enabled the application of Equation 22, which in turn represents a fundamentally novel way of addressing multiple toxic endpoints and/or alternative modeling assumptions in quantitative risk assessment.

The application of the proposed systematic probabilistic framework to a case involving TCE also illustrates several methods useful for addressing a number of technical issues necessary to consider in order to perform quantitative risk assessment for combined cancer and noncancer endpoints within a systematic probabilistic framework. For example, the quantitative analysis of interindividual variability in pharmacokinetic relations for TCE done in this case study involved a variety of methods that might also be useful in similar analyses for other compounds exhibiting saturable metabolism. Another example is the application of Equations 24a-b to calculate a dose-vs.-risk-of-adverse-response cdf that reflects uncertainty in the estimated parameters of an assumed dose-response function fit to available data. These

equations allow an analytic approach to this problem in the case of approximately normally distributed data, in contrast to Monte Carlo methods (e.g., as proposed by Slob and Pieters, 1998) that may not be practical to implement reliably, as shown in this case study (see Section 3.2).

All estimates of individual risk obtained in this study (Table 4) are far less than corresponding upper-bound point estimates of individual risk, $\hat{R}_{RME}$ and $\hat{R}_{High}$, that were obtained for comparison by Daniels et al. (2000) using standard, traditional deterministic methods (namely, algebraic substitution of upper-bound and/or default parameter values into equations used to estimate risk). Daniels et al. (2000), who did not consider JUV in pharmacokinetic and dose-response relations pertaining to TCE risk, obtained a similar result. In the present study, the systematic probabilistic framework used to consider JUV in pharmacokinetic and dose-response relations pertaining to TCE had a substantial impact on predicted risk for Site LF-13. This impact can be assessed by comparing the risk summary in Table 4 of the present study with that in Table 3 of Daniels et al. (2000). The upper-bound risk estimators $\overline{R}_{0.95}$ and $\langle R \rangle_{0.95}$ obtained in the present study are about 60- and 80-fold less, respectively, than the value of these estimators obtained by Daniels et al. (2000).

Even the JUV-estimator $R_{0.95,0.95}$ approximated in Daniels et al. (2000) is (slightly) greater than the value (of ~2×10^{-6}) obtained here using a nested Monte-Carlo procedure for the more conservative JUV-estimator $R_{0.95,0.99}$. Only the even more conservative JUV-estimator, $R_{0.99,0.99}$, obtained (37×10^{-6}) was substantially greater than 10^{-6}. While the latter value is less than the deterministically (hence, relatively easily) calculated $\hat{R}_{RME}$ value of (~60×10^{-6}), the two values are fairly close, indicating that $\hat{R}_{RME}$ in this case provides a credible estimate of the more precisely defined estimator $R_{0.99,0.99}$ (namely, the 99th percentile on uncertainty in risk to the person who is at the 99th percentile of risk relative to others at risk). Furthermore, $\hat{R}_{RME}$ is relatively easily calculated, whereas $R_{0.99,0.99}$ required about 10 h of computation in the present study. It might therefore be preferable to use $\hat{R}_{RME}$ to optimize risk reduction relative to an upper-bound JUV estimator (such as $R_{0.99,0.99}$).

Upper-bound JUV estimators allow explicit consideration of equity in the distribution of interindividual variability in imposed risk. Point estimates such as $\hat{R}_{RME}$ cannot do this explicitly, because they cannot generally be interpreted in any precise manner with respect to variability *per se* or to uncertainty *per se*. In the present case study, both $\hat{R}_{RME}$ and $R_{0.99,0.99}$ are $< 5 \times 10^{-4}$, which indicates (vaguely via $\hat{R}_{RME}$, explicitly by $R_{0.99,0.99}$) the *de minimis* nature of predicted upper-bound risks plausibly due to TCE at Site LF-13 hypothetically faced by those who would be among most at risk relative to others exposed to ground water from that site. However, such risk estimates do not necessarily correspond to the magnitude of health consequences predicted to be associated with such exposure. Such effects can *only* be addressed by quantitatively considering *uncertainty* in *population* risk, which in turn can only be accomplished by quantitative JUV analysis that characterizes uncertainty in population-average risk $\overline{R}$ conditional on population size n (e.g., via Equation 22) (Bogen, 1986, 1990b; Bogen and Spear, 1987).

Interesting results concerning population risk were obtained in this study via quantitative JUV analysis addressing multiple health endpoints and multiple mechanisms concerning TCE-induced health risk. Earlier results by Daniels et al. (2000) indicated that exposure to TCE from ground water at Site LF-13 would be *unlikely* to cause a single occurrence of a TCE-related health impact provided that $n < \sim 30,000$. Results from the present, more comprehensive analysis (Table 5) indicate that *a single case most likely would not occur even if the number of people served* by ground water from Site LF-13 *were ten million*. That is, the new results indicate that a single case is unlikely to occur under *any* realistic assumption concerning population size. Moreover, results obtained in the present study indicate that, under the assumptions used, there is a *>99.5% chance* that TCE-related mitigation of Site LF-13 would confer *no (i.e., zero) public-health benefit* if as many as 10,000 or fewer people were residentially served by the site's TCE-contaminated water (assuming the concentration of TCE never were to increase above levels measured in 1997). Therefore, under this population scenario, any resources directed at mitigating the site are *virtually certain* to be wasted from a public-health perspective. Even this hypothetical scenario is conservative, because it is likely that current TCE contamination in ground water at Site LF-13 (due to a finite mass of TCE contamination) could not persist for 70 y if that water were to serve 10,000

hypothetical residents throughout this period. Dilution of the source mass is expected, and the magnitude of this dilution is expected to be proportional to the water flow rate into and away from the source; indeed, this is the basis for pump-and-treat site-mitigation strategies. Water service to that many people from this single goundwater source—if even possible—would also induce some (perhaps substantial) infiltration of non-contaminated water, again causing dilution of residentially delivered TCE concentrations.

5. REFERENCES

Abbas, R., and J.W. Fisher. 1997. A physiologically based pharmacokinetic model for trichloroethylene and its metabolites, chloral hydrate, trichloroacetate, dichloroacetate, trichloroethanol, and trichloroethanol glucuronide in B6C3F1 mice. *Toxicol. Appl. Pharmacol.* 147:15-30.

Acharya, S., K. Mehta, S. Rodriguez, J. Pereira, S. Krishnan, and C.V. Rao. 1997. A histopathological study of liver and kidney in male Wistar rats treated with subtoxic doses of t-butyl alcohol and trichloroacetic acid. *Exper. Toxicol. Pathol.* 49:369-373.

Aitchison, J., and J.A.C. Brown. 1957. The Lognormal Distribution. Cambridge University Press, New York.

Allen, B.C., and J.W. Fisher. 1993. Pharmacokinetic modeling of trichloroethylene and trichloroacetic acid in humans. *Risk Anal.* 15:71-86.

Ames, B., and L.S. Gold. 1990a. Too many rodent carcinogens: Mitogenesis increases mutagenesis. *Science.* 249:970-972.

Ames, B.N., and L.S. Gold. 1990b. Chemical carcinogenesis: Too many rodent carcinogens. *Proc. Natl. Acad. Sci.* 87:7772-7776.

Ames, B.N., L.S. Gold, and W.C. Willett. 1995. The causes and prevention of cancer. *Proc. Natl. Acad. Sci.* 92:5258-5265.

Ames, B.N., M.K. Shigenaga, and L.S. Gold. 1993. DNA lesions, inducible DNA repair, and cell division: Three key factors in mutagenesis and carcinogenesis. *Environ. Health Perspect.* 101(suppl. 5):35-44.

Andersen, M.E., H.A. Barton, R. Bull, and I. Schultz. 1998. DCA dosimetry: Interpreting DCA-induced liver cancer dose-response and the potential for DCA to contribute to TCE-induced liver cancer, AL-OE-DR-TR-1998-0009. United States Air Force Armstrong Laboratory, Brooks Air Force Base, TX.

Anderson, E.L., R.E. Albert, R. McGaughy, L. Anderson, S. Bayard, D. Bayliss, C. Chen, M. Chu, H. Gibb, B. Haberman, C. Hiremath, D. Singh, and T. Thorslund. 1983. Quantitative approaches in use to assess cancer risk. *Risk Anal.* 3:277-295.

Armitage, P., and R. Doll. 1957. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br. J. Cancer.* 11:161-169.

Baird, S.J.S., J.T. Cohen, J.D. Graham, A.I. Shlyakter, and J.S. Evans. 1996. Noncancer risk assessment: A probabilistic alternative to current practice. *Human Ecol. Risk Assessment.* 2:79-102.

Bell, Z.G., K.J. Olsen, and T.J. Benya. 1978. Final Report of Audit Findings of the Manufacturing Chemists Association (MCA): Administered Trichloroethylene (TCE) Chronic Inhalation Study at Industrial Bio-Test Laboratories, Inc., Decator, Illinois. Unpublished study reported in USEPA (1985).

Bogen, K.T. 1986. *Uncertainty in Environmental Health Risk Assessment: A Framework for Analysis and an Application to a Chronic Exposure Situation Involving a Chemical Carcinogen.* Doctoral dissertation. University of California, Berkeley, School of Public Health. Berkeley, CA.

Bogen, K.T. 1988. Pharmacokinetics for regulatory risk assessment: The case of trichloroethlyene. *Regulatory Toxicol. Pharmacol.* 8:447-466.

Bogen, K.T. 1989. Cell proliferation kinetics and multistage cancer risk models. *J. National Cancer Inst.* 81:267-277.

Bogen, K.T. 1990a. Risk extrapolation for chlorinated methanes as promoters vs initiators of multistage carcinogenesis. *Fund. Appl. Toxicol.* 15:536-557.

Bogen, K.T. 1990b. Uncertainty in Environmental Health Risk Assessment. Garland Publishing, Inc., New York.

Bogen, K.T. 2000. *RiskQ 4.0: An Interactive Approach to Probability, Uncertainty and Statistics for use with Mathematica®.* UCRL-MA-110232 Rev. 1. Lawrence Livermore National Laboratory, Livermore, CA.

Bogen, K.T. 1994. Cancer potencies of heterocyclic amines found in cooked foods. *Fd. Chem. Toxicol.* 32:505-515.

Bogen, K.T. 1995. Methods to approximate joint uncertainty and variability in risk. *Risk Anal.* 15:411-419.

Bogen, K.T., and L.S. Gold. 1997. Trichloroethylene cancer risks: Simplified calculation of PBPK-based MCLs for cytotoxic endpoints. *Regulatory Toxicol. Pharmacol.* 25:26-42.

Bogen, K.T., and L.C. Hall. 1989. Pharmacokinetics for regulatory risk assessment: The case of 1,1,1-trichloroethane (Methyl chloroform). *Regulatory Toxicol. Pharmacol.* 10:26-50.

Bogen, K.T., L.C. Hall, T.E. McKone, D.W. Layton, and S.E. Patton. 1988. Health Risk Assessment of Trichloroethylene in California Drinking Water. Report prepared for the California Public Health Foundation and California Department of Health Services. UCRL-21007. Lawrence Livermore National Laboratory, Livermore, CA.

Bogen, K.T., and R.C. Spear. 1987. Integrating uncertainty and interindividual variability in environmental risk assessment. *Risk Anal.* 7:427-436.

Brown, L.P., D.G. Farrar, and C.G. DeRooij. 1990. Health risk assessment of environmental exposure to trichloroethylene. *Regulatory Toxicol. Pharmacol.* 11:24-41.

Bull, R.J., I.M. Sanchez, M.A. Nelson, J.L. Larson, and A.J. Lansing. 1990. Liver tumor induction in B6C3F1 mice by dichloroacetate and trichloroacetate. *Toxicology.* 63:341-359.

Bull, R.J. 2000. Mode of action of liver tumor induction by trichloroethylene and its metabolites, trichloroacetate and dichloroacetate. *Environ. Health Perspect.* 108(suppl. 2): 241-259.

Butterworth, B.E., and M.S. Bogdanffy. 1999. A comprehensive approach for integration of toxicity and cancer risk assessments. *Regulatory Toxicol. Pharmacol.* 29:23-36.

California Environmental Protection Agency (CalEPA). 1996. Air Toxics Hot Spots Program Risk Assessment Guidelines Part IV, Technical Support Document: Exposure Assessment and Stochastic Analysis (December 1996). CalEPA Office of Health Hazard Assessment, Berkeley, CA.

Carlson-Lynch, H., P.S. Price, J.C. Swartout, M.L. Dourson, and R.E. Keenan. 1999. Application of quantitative information on the uncertainty in the RfD of Noncarcinogenic risk assessments. *Human Ecol. Risk Assess.* 5:527-546.

Chaubey, Y.P., and G.S. Mudholkar. 1982. A new approximation for the distribution of the difference of two t-variables. *Commun. Statist.–Theor. Meth.* 11:2335-2342.

Cohen, S.M., and L.B. Ellwein. 1990. Cell proliferation in carcinogenesis. *Science.* 249:1007-1011.

Cohen, S.M., and L.B. Ellwein. 1991. Genetic errors, cell proliferation, and carcinogenesis. *Cancer Res.* 51:6493-6505.

Daniels, J., K.T. Bogen, and L. Hall. 1999. Procedures for Addressing Uncertainty and Variability in Exposure to Characterize Potential Health Risk From Trichloroethylene Contaminated Groundwater at Beale Air Force Base in California. UCID-CR-135784, Rev. 1. Lawrence Livermore National Laboratory, Livermore, CA.

Daniels, J., K.T. Bogen, and L. Hall. 2000. Analysis of uncertainty and variability in exposure to characterize risk: Case study involving trichloroethylene groundwater contamination at Beale Air Force Base in California. *Water, Air, and Soil Pollution* 123:273-298.

DeAngelo, A.B., F.B. Daniel, B.M. Most, and G.R. Olson. 1997. Failure of monochloroacetic acid and trichloroacetic acid administered in the drinking water to produce liver cancer in male F344/N rats. *J. Toxicol. Environ. Health.* 52:425-445.

DeAngelo, A.B., F.B. Daniel, J.A. Stober, and G.R. Olsen. 1991. The carcinogenicity of dichloroacetic acid in the male B6C3F1 mouse. *Fund. Appl. Toxicol.* 16:337-347.

Dees, C., and C. Travis. 1994. Trichloroacetate stimulation of liver DNA synthesis in male and female mice. *Toxicol. Lett.* 70:343-355.

Dourson, M.L., S.P. Felter, and D. Robinson. 1996. Evolution of science-based uncertainty factors in noncancer risk assessment. *Regulatory Toxicol. Pharmacol.* 24:108-120.

Eyre, R.J., D.K. Stevens, J.C. Parker, and R.J. Bull. 1995. Renal activation of trichloroethylene and S-(1,2-dichlorovinyl)-L-cysteine and cell proliferative responses in the kidneys of F344 rats and B6C3F1 mice. *J. Toxicol. Environ. Health.* 46:465-481.

Fahrig, R., S. Madle, and H. Baumann. 1995. Genetic toxicology of trichloroethylene. *Mutat. Res.* 340:1-36.

Finley, B., D. Proctor, P. Scott, N. Harrington, D. Paustenback, and P. Prince. 1994. Recommended distributions for exposure factors frequently used in health risk assessment. *Risk Anal.* 14:533-553.

Fisher, J.W., and B.C. Allen. 1993. Evaluating the risk of liver cancer in humans exposed to trichloroethylene using physiological models. *Risk Anal.* 15:87-95.

Fisher, J.W., M.L. Gargas, B.C. Allen, and M.E. Andersen. 1991. Physiologically based pharmacokinetic modeling with trichloroethylene and its metabolite, trichloroacetic acid, in the rat and the mouse. *Toxicol. Appl. Pharmacol.* 109:183-195.

Fisher, J.W., D. Mahle, and R. Abbas. 1998. A human physiologically based pharmacokinetic model for trichloroethylene and its metabolites, trichloroacetic acid and free trichloroethanol. *Toxicol. Appl. Pharmacol.* 152:339-359.

Gaylor, D.W., R.L. Kodell, J.J. Chen, and D. Krewski. 1999. A unified approach to risk assessment for cancer and noncancer endpoints based on benchmark doses. *Regulatory Toxicol. Pharmacol.* 29:151-157.

Ghosh, B.K. 1975. On the distribution of the difference of two *t*-variables. *J. Am. Statist. Assoc.* 70:463-467.

Harrington-Brock, K., C.L. Doerr, and M.M. Moore. 1998. Mutagenicity of three disinfection by-products: Di- and trichloroacetic acid and chloral hydrate in L5178Y/TK +/- (-)3.7.2C mouse lymphoma cells. *Mutat. Res.* 413:265-276.

Herren-Freund, S.L., M.A. Pereira, M.D. Khoury, and G. Olson. 1987. The carcinogenicity of trichloroethylene and its metabolites, trichloroacetic and dichloroacetic acid, in mouse liver. *Toxicol. Appl. Pharmacol.* 90:183-189.

Iman, R.L., and W.J. Conover. 1982. A distribution-free approach to inducing rank correlation among input variates. *Commun. Statist. (Ser. B) Simulat. Computat.* 11:311-334.

Jennrich, R.I. 1970. An asymptotic χ^2 test for the equality of two correlation matrices. *J. Am. Statist. Assoc.* 65:904-912.

Larson, J.L., and R.J. Bull. 1992a. Metabolism and lipoperoxidative activity of trichloroacetate and dichloroacetate in rats and mice. *Toxicol. Appl. Pharmacol.* 115:268-277.

Larson, J.L., and R.J. Bull. 1992b. Species differences in the metabolism of trichloroethylene to the carcinogenic metabolites trichloroacetate and dichloroacetate. *Toxicol. Appl. Pharmacol.* 115:278-285.

Lash, L.H., J.W. Fisher, J.C. Lipscomb, and J.C. Parker. 2000a. Metabolism of trichloroethylene. *Environ. Health Perspect.* 108(suppl. 2):177-200.

Lash, L.H., J.C. Parker, and C.S. Scott. 2000b. Modes of action of trichloroethylene for kidney tumorigenesis. *Environ. Health Perspect.* 108(suppl. 2):225-240.

Lewis, S.C. 1993. Reducing uncertainty with adjustment factors: Improvements in quantitative noncancer risk assessment. *Fund. Appl. Toxicol.* 20:2-4.

Lipscomb, J.C., J.W. Fisher, P.D. Confer, and J.Z. Byczkowski. 1998. *In vitro* to *in vivo* extrapolation for trichloroethylene metabolism in humans. *Toxicol. Appl. Pharmacol.* 152:376-387.

Maltoni, C., G. Lefemine, and G. Cotti. 1986. Archives of Research on Industrial Carcinogenesis. Vol. V. Experimental Research of Trichloroethylene Carcinogenesis. Princeton University Press, Princeton, NJ.

Merdink, J.L., A. Gonzalez-Leon, R.J. Bull, and I.R. Schultz. 1998. The extent of dichloroacetate formation from trichloroethylene, chloral hydrate, trichloroacetate, and trichloroethanol in B6C3F1 mice. *Toxicol. Sci.* 45:33-41.

Moolgavkar, S.H. 1983. Model for human carcinogenesis: Action of environmental agents. *Environ. Health Perspect.* 50:285-291.

Moolgavkar, S.H., A. Dewanji, and D.J. Venzon. 1988. A stochastic two-stage model for cancer risk assessment: The hazard function and the probability of tumor. *Risk Anal.* 8:383-392.

Moolgavkar, S.H., and A.G. Knudson. 1981. Mutation and cancer: A model for human carcinogenesis. *J. Natl. Cancer Inst.* 66:1037-1052.

Moore, M.M., and K. Harrington-Brock. 2000. Mutagenicity of trichloroethylene and its metabolites: implications for the risk assessment of trichloroethylene. *Environ. Health Perspect.* 108(suppl. 2):215-223.

National Cancer Institute (NCI). 1976. Carcinogenesis Bioassay of Trichloroethylene. NCI-CG-TR-2, DHEW Publ. No. (NIH) 76-802. U.S. Government Printing Office, Washington, DC.

National Research Council (NRC). 1994. *Science and Judgment in Risk Assessment.* NRC Committee on Risk Assessment of Hazardous Air Pollutants, National Academy Press, Washington, DC.

National Toxicology Program (NTP). 1988. *Toxicology and Carcinogenesis Studies of Trichloroethylene (CAS No. 79-01-6) in Four Strains of Rats (ACI, August, Marshall, Osborne-Mendel) (Gavage Studies).* NIH Pub No. 88-2529, NTP Tech. Rep. Ser. No. 273. National Institutes of Health, NTP, Research Triangle Park, NC.

National Toxicology Program (NTP). 1990. *Carcinogenesis Studies of Trichloroethylene (Without Epichlorohydrin) (CAS No. 79-01-6) in F344/N Rates and B3C3F1 Mice (Gavage Studies).* NIH Pub No. 90-1799. National Institutes of Health, NTP, Research Triangle Park, NC.

Ni, Y.C., T.Y. Wong, R.V. Llyoyd, T.M. Heinze, S. Shelton, D. Caciano, F.F. Kadlubar, and P.P. Fu. 1996. Mouse liver microsomal metabolism of chloral hydrate, trichloroacetic acid, and trichloroethanol leading to induction of lipid peroxidation via a free radical mechanism. *Drug Metab. Disposition* 24:81-90.

Ojo, M.O. 1988. Approximation to the distribution of the difference of two weighted and unweighted t-variables. *J. Statist. Res. (Bangladesh)* 22:37-42.

Patil., V.H. 1965. Approximation to the Behrens-Fisher distributions. *Biometrika* 52:267-271.

Pereira, M.A. 1996. Carcinogenic activity of dichloroacetic acid and trichloroacetic acid in the liver of female B6C3F1 mice. *Fund. Appl. Toxicol.* 31:192-199.

Pereira, M.A., and J.B. Phelps. 1996. Promotion by dichloroacetic acid and trichloroacetic acid of N-methyl-N-nitrosourea-initiated cancer in the liver of female B6C3F1 mice. *Cancer Lett.* 102:133-141.

Purrier. W. 1997. Personal communication regarding detailed data for TCE groundwater concentration at Site LF-13 of Beale Air Force Base, California. Law Engineering and Environmental Services, Inc., Sacramento, CA (December 1997).

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. 1992. Numerical Recipes in FORTRAN — The Art of Scientific Computing. Cambridge University Press, New York, NY, pp. 650-700.

Renwick, A.G. 1993. Data derived safety factors for the evaluation of food additives and environmental contaminants. *Food Add. Contam.* 10:275-305.

Rubin, H. 1960. On the distribution of the weighted difference of two independent Student variables. *J. Roy. Statist. Soc., Ser. B* 22:188-194.

Singh, K.P. 1990. Computing probabilities for the difference of two t-variables. In: Page, C., and R. LePage, eds. *Computing Science and Statistics, Statistics of Many Parameters: Curves, Images, Spatial Models.* Proceedings of the 22nd Symposium on the Interface, East Lansing, MI, May 16-19, 1990. Springer-Verlag, New York, NY, pp. 519-523.

Slob, W., and M.N. Pieters. 1998. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. *Risk Anal.* 18:787-798.

Stenner, R.D., J.L. Merdink, J.W. Fisher, and R.J. Bull. 1998. Physiologically-based pharmacokinetic model for trichloroethylene considering enterohepatic recirculation of major metabolites. *Risk Anal.* 18:261-269.

Templin, M.V., J.C. Parker, and R.J. Bull. 1993. Relative formation of dichloroacetate and trichloroacetate from trichloroethylene in male B6C3F1 mice. *Toxicol. Appl. Pharmacol.* 123:1-8.

U.S. Environmental Protection Agency (USEPA). 1985. *Health Assessment Document for Trichloroethylene.* EPA/600/8-82/006F. USEPA Office of Research and Development, Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

U.S. Environmental Protection Agency (USEPA). 1987a. Technical Analysis of New Methods and Data Regarding Dichloromethane Hazard Assessments. EPA/600/8-87/029A (June 1987). USEPA Office of Health and Environmental Assessment, Washington, DC.

U.S. Environmental Protection Agency (USEPA). 1987b. Addendum to the Health Assessment Document for Trichloroethylene: Updated Carcinogenicity Assessment for Trichloroethylene. EPA/600/8-82/006FA. USEPA Office of Research and Development, Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

U.S. Environmental Protection Agency (USEPA). 1992. Draft report: A cross-species scaling factor for carcinogen risk assessment based on equivalence of $mg/kg^{3/4}/day$. *Fed. Register.* 57(No. 109, June 5):24152-24172.

U.S. Environmental Protection Agency (USEPA). 1996. Proposed Guidelines for Carcinogen Risk Assessment. EPA/600/P-92/003C. USEPA Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). 1998. Draft Water Quality Criteria Methodology: Human Health. Federal Register Notice. [*Fed. Regist.* 63(No. 157, Aug. 14):43755-43828] EPA/822-Z-98-001. USEPA Office of Water, Washington, DC.

URS Greiner Woodward Clyde (URSGWC). 1998. Management Action Plan, Beale Air Force Base, California. December 1998. Prepared for Headquarters Air Combat Command (ACC), Langley Air Force Base, VA, under contract to US Air Force Center for Environmental Excellence (AFCEE), Brooks Air Force Base, TX, Project No. ACCH19987544 [obtain from Chief, Environmental Restoration (M.E. O'Brien), Beale Air Force Base, California]. URSGWC, Omaha, NE.

Walker, G.A., and J.G. Saw. 1978. The distribution of linear combinations of t-variables. *J. Am. Statist. Assoc.* 73:876-878.

Weil, C.S. 1972. Statistics vs. safety factors and scientific judgment in the evaluation of safety for man. *Toxicol. Appl. Pharmacol.* 21:454-463.

Wolfram, S. 1999. *The Mathematica Book.* Cambridge University Press, Cambridge, UK.

Wright, S. 1992. Adjusted p-values for simultaneous inference. *Biometrics.* 48:1005-1013.

Appendix 1

Method of Moments for Lognormal Variates

Given a normally distributed variate Y with arithmetic mean (AM) μ_Y, standard deviation (SD) σ_Y, and corresponding coefficient of variation (CV = SD/AM) $\gamma_Y = (\sigma_Y / \mu_Y)$, the variate $X = e^Y$ has a lognormal (LN) distribution with geometric mean (GM) e^{μ_Y} and geometric standard deviation (GSD) e^{σ_Y}, where $e = \ln^{-1}(1)$ and ln denotes natural logarithm. These assumptions are efficiently denoted $Y \sim N(\mu_Y, \sigma_Y)$ and $X \sim LN(\mu_Y, \sigma_Y)$. The method of moments may be used to relate given AM, SD, and CV values of X (μ_X, σ_X, and γ_X, respectively) to those of Y; in particular, the AM/GM ratio for X, $\rho = (\mu_X / e^{\mu_Y})$, is equal to $e^{\sigma_Y^2/2}$ where $\sigma_Y^2 = \ln \rho^2 = \ln(1 + \gamma_X^2)$ (Aitchison and Brown, 1957).

LN moment relations conveniently imply that the ratio of any given percentile of X relative to its GM or AM corresponds to a unique set of LN parameters. Let $X_p = e^{\mu_Y + \sigma_Y z_p}$ denote the $100p$th percentile of X, where $0 \leq p \leq 1$, $z_p = \Phi^{-1}(p)$, and Φ is the cumulative normal probability distribution function. Now let $q_p = X_p / e^{\mu_Y}$ and $r_p = X_p / \mu_X$ denote the ratios of X_p to the GM and AM of X, respectively. Conditional on e^{μ_Y} and the ratio q_p, it follows immediately (by solving for μ_Y and σ_Y) that $X \sim LN(\mu_Y, \ln q_p^{1/z_p})$. Conditional on μ_X and the ratio r_p, it follows that $X \sim LN[\ln(\mu_X) - (\sigma_Y^2/2), \sigma_Y]$, where σ_Y is the positive σ_Y-root of $\sigma_Y^2 - 2z_p\sigma_Y + 2\ln r_p = 0$; i.e., $\sigma_Y = z_p + \sqrt{z_p^2 - \ln r_p^2}$ for all $r_p \leq e^{z_p^2}/2$ (larger values of r_p are not possible conditional on z_p).

LN moment relations also imply that for any independent LN variates $X_i = e^{Y_i}$, $i = 1,...,n$, with corresponding CVs γ_i, the CV γ_Z of the product $Z = \prod_{i=1}^{n} X_i$ is conveniently ralated as follows to the CVs γ_i of X_i:

$$\gamma_Z = \sqrt{e^{\sigma_{\ln Z}^2} - 1} = \sqrt{-1 + \exp\left(\sigma_{\Sigma_{i=1}^n Y_i}^2\right)}$$

$$= \sqrt{-1 + \exp\left(\sum_{i=1}^{n} \sigma_{Y_i}^2\right)} = \sqrt{-1 + \exp\left(\sum_{i=1}^{n} \ln\left(1 + \gamma_i^2\right)\right)}$$

$$= \sqrt{-1 + \prod_{i=1}^{n}\left(1 + \gamma_i^2\right)} \quad .$$

Conditional on known γ_Z and γ_i for $i \neq j$ and $1 \leq j \leq n$, inverting the latter equation readily yields the unknown CV of X_j as

$$\gamma_j = \sqrt{\frac{1 + \gamma_Z^2}{\prod_{i \neq j}^{n}\left(1 + \gamma_i^2\right)} - 1} \quad .$$

If γ_j is obtained in this way and a single additional X_j-parameter among the set {AM, SD, GM, GSD} is known, all three remaining X_j-parameters are easily obtained via the moment relations described above. For example, if the AM of X_j equals 1, it follows that $X_j \sim LN(-\sigma^2/2, \sigma)$ where $\sigma^2 = \ln(1 + \gamma_j^2)$.

Appendix 2

The Exact Distribution of a
Linear Function of Correlated t-Distributed Variates

Exact and corresponding approximate expressions are available for the distribution of linear functions of independent t-distributed variates (Ruben, 1960; Patil, 1965; Ghosh, 1975; Walker and Saw, 1978; Chaubey and Mudholkar, 1982; Ojo, 1988; Singh, 1990), but not for correlated t-distributed variates. The exact distribution of a particular weighted difference (or sum) of two independent (uncorrelated) t-distributed variates, known as the Behrens-Fisher distribution, was derived by Ruben (1960) in integral form. A linear combination of t-distributed variates is referred to as a generalized Behrens-Fisher distribution (Patil, 1965). The result obtained by Ruben (1960) is thus one type of generalized Behrens-Fisher distribution. However, his approach can be generalized as follows to obtain the exact distribution of any linear combination of t-variates with specified correlations.

Let $S = \sum_{i=1}^{n} w_i T_{f_i}$ be a general linear function consisting of a weighted sum of correlated, t-distributed variates T_{f_i}, each with corresponding weights w_i and degrees of freedom f_i for $i = 1,...,n$. Now introduce two corresponding sets of auxiliary variates Z_i and U_i $(i = 1,...,n)$, where Z_i are correlated $N(0,1)$ (standard normal) variates all with zero mean, unit variance, and an ($n{\times}n$ symmetric) correlation matrix $\mathbf{R} = (\rho_{i,j})$ ($\{i,j\} = 1,...,n$), and where U_i have independent chi-square distributions with f_i corresponding degrees of freedom. Below, $g_X(x)$ and $G_X(x)$ denote the density function and cumulative distribution function (cdf), respectively, of the specified variate (X). It follows that

$$S = \sum_{i=1}^{n} \frac{w_i Z_i}{\sqrt{U_i/f_i}} , \qquad (2.1)$$

so that, conditional on variates $\mathbf{U} = (U_1, ..., U_n)$, S is distributed as $N(0,\sigma_S^2)$, with zero mean, with variance

$$\sigma_S^2 = \sum_{i=1}^{n} \frac{w_i^2 f_i}{U_i} + 2\sum_{i<j} \rho_{i,j} w_i w_j \sqrt{\frac{f_i f_j}{U_i U_j}} , \qquad (2.2)$$

72

and with the corresponding density function,

$$g_S(s \mid \mathbf{U}) = \frac{1}{\sqrt{2\pi}\,\sigma_S}\exp\!\left(\frac{-s^2}{2\sigma_S^2}\right). \tag{2.3}$$

The unconditional density function of S is then given by

$$g_S(s) = \int_0^\infty \!\!\cdots\! \int_0^\infty g_S(s|\mathbf{u})\prod_{i=1}^n \frac{e^{-\frac{1}{2}u_i}\,u_i^{\left(\frac{f_i}{2}-1\right)}}{2^{\frac{1}{2}f_i}\,\Gamma\!\left(\frac{1}{2}f_i\right)}\,du_i\ , \tag{2.4}$$

Now transform u_i in terms of new variates $y = \sum_{i=1}^n u_i$ and $x_i = u_i/y$ for $i = 1,\ldots,n-1$, and also define $x_n = 1 - \sum_{i=1}^{n-1} x_i$ and $\phi^2(\mathbf{x}) = f/(y\sigma_S^2)$, where $\mathbf{x} = (x_1, \ldots, x_{n-1})$ and $f = \sum_{i=1}^n f_i$. This transformation implies that

$$\phi^2(\mathbf{x}) = \frac{f}{y\,\sigma_S^2} = f\!\left(\sum_{i=1}^n \frac{w_i^2 f_i}{x_i} + 2\sum_{i<j}\rho_{i,j}\,w_i\,w_j\sqrt{\frac{f_i f_j}{x_i x_j}}\right)^{-1}, \tag{2.5}$$

and it has the Jacobian matrix: $\mathbf{J} = (y\mathbf{I}_n + \mathbf{M})$, where $\mathbf{I}_n$ is an $n{\times}n$ identity matrix and the elements $m_{i,j}$ of the $n{\times}n$ matrix $\mathbf{M}$ are all zero, except $m_{i,n} = -y$ for $i = 1,\ldots,n-1$ and $m_{n,j} = x_j$ for $j = 1,\ldots,n$. The corresponding Jacobian, $|\mathbf{J}| = \text{Det}(\mathbf{J}) = y^{n-1}$, and Equation 2.5, reduce Equation 2.4 to

$$g_S(s) = \int_0^1 \!\!\cdots\! \int_0^1 \left[\int_0^\infty g_S(s|\mathbf{u})\,y^{n-1}\,e^{\left(\frac{-y}{2}\right)}\,y^{\left(\frac{f}{2}-n\right)}dy\right]2^{\left(\frac{-f}{2}\right)}\prod_{i=1}^n \frac{x_i^{\left(\frac{f_i}{2}-1\right)}}{\Gamma\!\left(\frac{1}{2}f_i\right)}\,dx_1\ldots dx_{n-1}\ , \tag{2.6a}$$

$$= \int_0^1 \!\!\cdots\! \int_0^1 \left[\int_0^\infty e^{-\frac{y}{2}\left(1+\frac{s^2\phi^2(\mathbf{x})}{f}\right)}\,y^{\left(\frac{f}{2}-1\right)}dy\right]\frac{\phi(\mathbf{x})}{\sqrt{f}}2^{\left(\frac{-f}{2}\right)}\prod_{i=1}^n \frac{x_i^{\left(\frac{f_i}{2}-1\right)}}{\Gamma\!\left(\frac{1}{2}f_i\right)}\,dx_1\ldots dx_{n-1} \tag{2.6b}$$

$$= \int_0^1 \!\!\cdots\! \int_0^1 \left\{\frac{\left[1 + \left(s^2\phi^2(\mathbf{x})/f\right)\right]^{\frac{1}{2}(f-1)}}{\sqrt{f}\ \text{B}\!\left(\frac{f}{2},\frac{1}{2}\right)}\,\phi(\mathbf{x})\right\}\Gamma\!\left(\frac{f}{2}\right)\prod_{i=1}^n \frac{x_i^{\left(\frac{f_i}{2}-1\right)}}{\Gamma\!\left(\frac{1}{2}f_i\right)}\,dx_1\ldots dx_{n-1}\ , \tag{2.6c}$$

in which $\phi(\mathbf{x})$ denotes the positive square root of $\phi^2(\mathbf{x})$ defined by Equation 2.5. A new transformation, $t' = s\,\phi(\mathbf{x})$ with Jacobian $J = 1/\phi(\mathbf{x})$, shows the braced expression in

Equation 2.6c to be J times the density $g_{T_f}(t')$ of a t-variate with f degrees of freedom (i.e., T_j) conditional on x. Integrating Equation 2.6c with respect to t' over the range $[-\infty, t]$ shows the cdf of S to involve a multiple integral with respect to x over the cdf for T_j evaluated at x-dependent arguments defined using Equation 2.5:

$$G_S(t) = \int_0^1 \cdots \int_0^1 G_{T_f}(t\,\phi(\mathbf{x})\,|\,\mathbf{x})\,\Gamma\left(\tfrac{f}{2}\right)\prod_{i=1}^{n} \frac{x_i^{\left(\frac{f_i}{2}-1\right)}}{\Gamma\left(\frac{1}{2}f_i\right)}\,dx_1...dx_{n-1}\quad.$$ (2.7)

In the case that $n = 2$, using x to denote x_1, Equation 2.7 reduces to

$$G_S(t) = \int_0^1 G_{T_f}(t\,\phi(x)\,|\,x)\,\frac{x^{\frac{f_1}{2}-1}(1-x)^{\frac{f_2}{2}-1}}{B\left(\frac{f_1}{2},\frac{f_2}{2}\right)}\,dx$$ (2.8a)

$$= \int_0^1 G_{T_f}(t\,\phi(x)\,|\,x)\left(dB\left(\tfrac{f_1}{2},\tfrac{f_2}{2},x\right)/dx\right)dx\quad,$$ (2.8b)

where $B(a,b)$ and $B(a,b,x)$ denote the beta function, and the corresponding incomplete beta-function ratio with respect to x, respectively, of the specified arguments (see Abbreviations and Notation). Thus, in this case, $G_S(t)$ is for given t the mean value of the conditional Student-t cdf, $G_{T_j}(t\,\varphi(x)\,|\,x)$, where x is Beta-distributed with parameters $f_1/2$ and $f_2/2$. The latter fact was noted by Ruben (1960), who obtained $G_S(t)$ specifically for the case in which $n = 2$, $w_1 = \sin\theta$, $w_2 = \cos\theta$, and $\rho_{1,2} = 0$.

74

Appendix 3

Documentation of *Mathematica* 4.0® and *RiskQ* Calculations

All calculations were performed on a 400-MHz PowerMac G4 using the programs *Mathematica*® 4.0 (Wolfram, 1999) and *RiskQ* 4.0 (Bogen, 2000). Documentation of these calculations appears in Appendices 3.A through 3.I which follow, in which calculations and related comments are organized by topic. Appendices 3.A (Concentration), 3.B (Intakes), and 3.C (Fraction of Lifetime at One Local Residence) all document the derivation or re-derivation of exposure-related input variates explained in Daniels et al. (2000), which were used to calculate TCE exposures as explained above (Section 2.2). Appendices 3.D (Effective Genotoxic Dose) and 3.E (Effective Cytotoxic Dose) document the calculation of corresponding biologically effective (TCE or TCA) doses. Note that calculations pertaining to the definition or characterization of variates V_w, V_{Vmax}, $V_{fm,ing}$, V_{fd}, $(f_{deq}/V_{t,P})$ and V_e all appear in Appendix 3.E. Appendix 3.F (Effective Dose Correlations) documents calculations made to estimate rank correlations among MA- and pathway-specific normalized biologically effective doses. Appendix 3.G (Potency) documents all calculations made pertaining to modeled dose-response under both mechanisms of carcinogenic action considered (MA$_G$ and MA$_C$). Appendix 3.H (TCE Risk) documents all calculations made pertaining to corresponding predicted risk. Note that calculations pertaining to the definition of variates U_{chron}, U_{tdyn}, and V_{tdyn} appear in Appendix 3.H. Finally, Appendix 3.I (Functions Used) briefly describes all *Mathematica*® and *RiskQ* 4.0 functions used for calculations documented in Appendices 3.A-2.H.

Please note that more detailed explanation of *Mathematica*®, *RiskQ*, and JUV analysis is beyond the scope of this report, and is provided in references cited.

Appendix 3.A

Concentration

of TCE @ BAFB (mg/L)

$Y = \text{Log}[X]$

```
con = {.018, .021, .028};  (*mg/L*)
{{my, sdmy} = {EV[Log[con]], SD[Log[con]] / Sqrt[3]},
{mx, sdmx} = {EV[con], SD[con] / Sqrt[3]}}
```

```
{{-3.81872, 0.129473}, {0.0223333, 0.00296273}}
```

```
{cv = sdmx/mx, E^sdmx}
```

```
{0.13266, 1.00297}
```

```
t0 = tt /. Solve[mx + tt * sdx == 0, tt][[1]]
```

```
-4.35212
```

```
RQ[C, T, 2, t0]
```

```
0.0244757
```

```
tsim = SimulateCdf[{T, 2}, 2000];
```

```
Sort[tsim][[{1, 2, 3, 1998, 1990, 1999, 2000}]]
```

```
{-31.607, -22.3327, -18.2209, 18.2209, 9.45819, 22.3327, 31.607}
```

```
c = E^(0 + sdmy*tsim);
{{mc = EV[c], SD[c]}, {EV[Log[c]], SD[Log[c]]}, Idf[Cdf[c], {.5, .95}]}
```

```
{{1.08122, 1.43409}, {1.10724×10^-17, 0.32201}, {0.999908, 1.45641}}
```

simulated conc. values

```
scon = (mx / mc) c;
{{mcon = EV[scon], SD[scon]}, {EV[Log[scon]], SD[Log[scon]]}, Idf[cdf = Cdf[scon], {.5, .95}]}
```

```
{{0.0223333, 0.0296222}, {-3.87976, 0.32201}, {0.0206538, 0.0300832}}
```

```
Take[Sort[scon], -10]
```

```
{0.074688, 0.0801479, 0.0871134, 0.0963267,
 0.109112, 0.128074, 0.159096, 0.218571, 0.372219, 1.23676}
```

```
PlotCdf[cdf, Xmin -> -.001, Xmax -> .05];
```
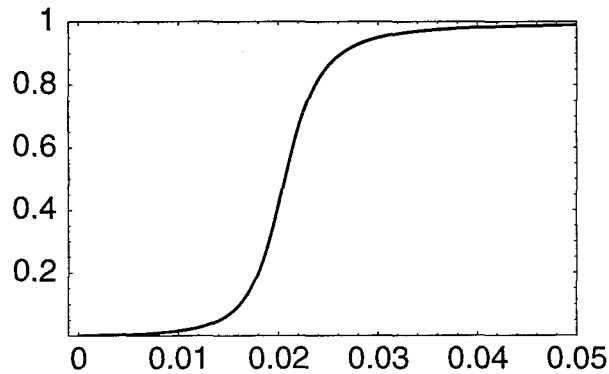


```
scdf = StandardizeCdf[cdf, 404];

WriteMatrix["BogenHD:Desktop Folder:concentration.txt", N[scdf]];
```

■ Log-Transform Utility Functions (where X = LogY)

```
MSDx[GMx_, GSDx_] := Module[{mux, sigy},
  sigy = Log[GSDx];
  mux = GMx E^((sigy^2) / 2);
  mux {1, Sqrt[E^(sigy^2) - 1]}]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
  sigy = Sqrt[Log[1 + (SDx / Mx)^2]];
  muy = Log[Mx] - (sigy^2) / 2;
  E^{muy, sigy}]
```

end

Appendix 3.B

Intakes

2-17-99 (updated 4-26-99)

```
<< RiskQ`;
```

■ Log-Transform Utility Functions

```
GMGSDx::usage =
   "GMGSDx[Mx,SDx] returns the  geometric mean and geometric sdandard deviation of
     a lognormal variate X that also has the specified arithmetic mean Mx and
     arithmetic sdandard deviation SDx, based on the method of moments.";

MSDx::usage =
   "MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation
     of a lognormal variate X that also has the specified geometric mean GMx and
     geometric sdandard deviation GSDx, based on the method of moments.";

GMGSDx1::usage =
   "GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that
     the product X1*X2 has the desired coefficient of variation (CV) = cvWant,
     conditional on the lognormal variate X2 having an arithmetic mean and
     CV equal to 1 and cv2, respectively, based on the method of moments.";

MSDx[GMx_, GSDx_] := Module[{mux, sigy},
        sigy = Log[GSDx];
        mux = GMxE^((sigy^2) / 2);
        mux{1, Sqrt[E^(sigy^2) - 1]}]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
        sigy = Sqrt[Log[1 + (SDx/ Mx) ^2]];
        muy = Log[Mx] - (sigy^2) / 2;
        E^{muy, sigy}]

GMGSDx1[cvWant_, cv2_] := Module[{my1},
```

$$muy1 = Log\left[Sqrt\left[\frac{cv2^2 + 1}{cvWant^2 + 1}\right]\right];$$

```
        E^{muy1, Sqrt[-2 muy1]}
   ] /; cvWant >= cv2
```

■ Data on 1998 U.S. Population

```
dat =
  Partition[{5, 18983, 10, 19928, 15, 19268, 20, 19535, 25,  17768, 30, 18545, 35, 20014,
      40, 22602, 45, 21962, 50, 18978, 55, 15907, 60, 12587, 65, 10332, 70, 9530, 75, 8782,
    80, 7227, 85, 4739, 90, 2554, 95, 1105, 100, 322, 105, 64}, 2];
TBL[data = Data[dat, {Age, temp}, Append -> {Pop, 1000 * temp},
    Drop -> temp, Append -> {{Fpop, 1. Pop / (Plus @@ Pop), 1},
      {CFpop, SUM[Fpop], 1}}]]
```

Age	Pop	Fpop	CFpop
5	18983000	0.0701173	0.0701173
10	19928000	0.0736078	0.143725
15	19268000	0.07117	0.214895
20	19535000	0.0721562	0.287051
25	17768000	0.0656295	0.352681
30	18545000	0.0684995	0.42118
35	20014000	0.0739255	0.495106
40	22602000	0.0834848	0.578591
45	21962000	0.0811208	0.659711
50	18978000	0.0700988	0.72981
55	15907000	0.0587555	0.788566
60	12587000	0.0464925	0.835058
65	10332000	0.0381632	0.873221
70	9530000	0.0352009	0.908422
75	8782000	0.032438	0.94086
80	7227000	0.0266943	0.967555
85	4739000	0.0175044	0.985059
90	2554000	0.00943368	0.994493
95	1105000	0.00408153	0.998574
100	322000	0.00118937	0.999764
105	64000	0.000236396	1.

```
tpop = Plus @@ Data[data, Pop]
```

270732000

12 years and over: 200899000

```
{plo, phi} = Data[data, CFpop][[{2, 3}]];
pwant = plo + (phi - plo) 2 / 5;
{plo, pwant, phi}
```

{0.143725, 0.172193, 0.214895}

```
{f12 = pwant, f12p = 1 - f12, f12 + f12p}
```

{0.172193, 0.827807, 1.}

18 years and over: 200899000

```
f2 = 2 Data[data, CFpop][[1]] / 5;
f18p = 200899000. / tpop;
f18 = 1 - f2 - f18p;
{w1, w2, w3, ww} = {f2, f18, f18p, f2 + f18 + f18p}
```

{0.0280469, 0.229895, 0.742059, 1.}

■ Ingestion (L/kg-d) Ershow and Cantor, 1989, Table 36 p. 76

```
age = {1, 10, 20, 65, 65 plus};
w = {87.7, 1127.2, 1197.8, 3960.7, 697.0} / 7070.4;
ingest = {53.2, 38.7, 18.4, 21.4, 23.1} / 1000;
sd = {50.9, 23.8, 10.7, 12.2, 9.7} / 1000;
gmgsd = MapThread[GMGSDx[#1, #2]&, {ingest, sd}]
```

{{0.0384398, 2.23934}, {0.032965, 1.76188}, {0.0159061, 1.71553}, {0.0185911, 1.69976},
  {0.0212984, 1.49628}}

```
Plus @@ w
```

1.

```
cdfs = LogNormalCdf[#[[1]], #[[2]], 1000] & /@ Log[gmgsd];
```

```
adf = AverageCdf[cdfs, Weights -> w];
Dimensions[adf]
```

{5001, 2}

```
sim = SimulateCdf[cdfs, 5000, Report -> False];
cdfTWA = Cdf[Plus @@ (w * sim)];
```

```
cdfTWAcorrected = Cdf[Plus @@ ({1, 10, 9, 45, 5} * sim / 70)];
```

```
{gm, gsd} = GMGSDx[24.2 / 1000, 17 / 1000];
{{gm, gsd}, Log[{gm, gsd}], 17 / 24.2}
```

{{0.0198023, 1.88387}, {-3.92196, 0.63333}, 0.702479}

```
cdfErCan = LogNormalCdf[-3.92195619921997042`, 0.633330172027946325`, 200];
```

```
PlotCdf[{cdfTWA, adf, cdfErCan}, Xmin → -.0001, Xmax → .06];
```
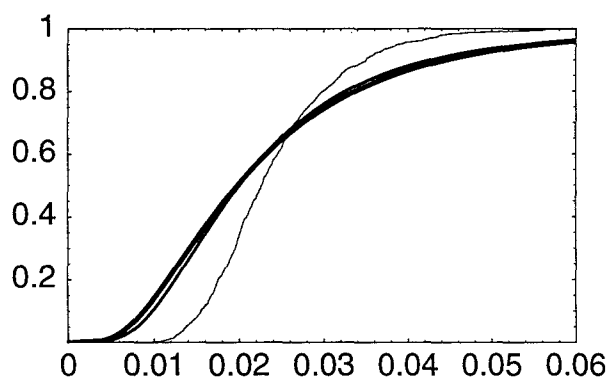
```
PlotCdf[{cdfTWA, cdfTWAcorrected}, Xmin → -.0001, Xmax → .06];
```
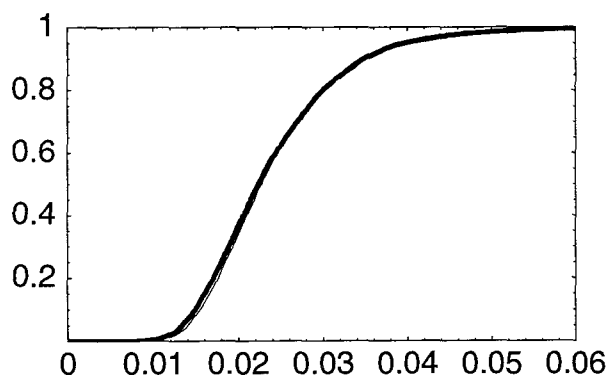


```
{EV[#], SD[#], Idf[#, {.5, .95}]}&/@{cdfTWA, cdfTWAcorrected, adf, cdfErCan}
```

```
{{0.0241953, 0.00802976, {0.0227623, 0.0390576}}, {0.0240441, 0.00862608,
    {0.0223285, 0.0398658}}, {0.024215, 0.0170167, {0.0200231, 0.054078}},
    {0.0242229, 0.0170924, {0.0198023, 0.0561224}}}
```

end

■ SABW Ratio calculations

■ Distribution of Body Surface Area to Body Weight (cm2/kg) Ratio (Phillips et al., 1993)

```
pp = {0, 5, 10, 25, 50, 75, 90, 95, 100} / 100.;
ratio = {
  {421, 470, 507, 563, 617, 719, 784, 846, 1142},
   {268, 291, 328, 376, 422, 454, 501, 594, 670},
   {200, 238, 244, 270, 286, 302, 316, 329, 351}} *1.;
cdfs = Transpose[{#, pp}]&/@ratio;
PlotCdf[cdfs];
```
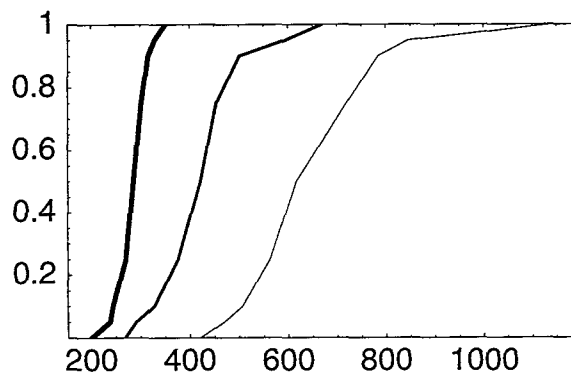
■ sabwALL

```
adf = AverageCdf[cdfs, Weights -> {w1, w2, w3}];
PlotCdf[adf, Xmin -> -.01, Xmax -> 800];
```
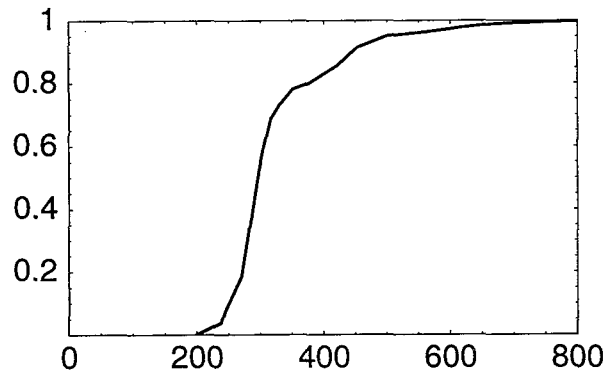


```
TBL[adf]
```

```
200.    0
238.    0.0371029
244.    0.0742059
268.    0.176952
270.    0.186514
286.    0.380025
291.    0.440497
302.    0.571456
316.    0.687114
328.    0.725091
329.    0.728664
351.    0.781572
376.    0.799532
421.    0.855756
422.    0.857034
454.    0.915424
470.    0.927621
501.    0.951541
507.    0.95251
563.    0.963639
594.    0.971495
617.    0.97796
670.    0.98962
719.    0.992988
784.    0.997195
846.    0.998598
1142.   1.
```

```
{EV[adf], Idf[adf, .95]}
```

```
{325.881, 499.003}
```

```
WriteMatrix["BogenHD:Desktop Folder:sabwratioALL.txt", adf];
```

```
{x, y} = Transpose[Rest[Drop[adf, -1]]]
```

```
{{238., 244., 268., 270., 286., 291., 302., 316., 328., 329., 351., 376.,
   421., 422., 454., 470., 501., 507., 563., 594., 617., 670., 719., 784., 846.},
  {0.0371029, 0.0742059, 0.176952, 0.186514, 0.380025, 0.440497, 0.571456, 0.687114,
   0.725091, 0.728664, 0.781572, 0.799532, 0.855756, 0.857034, 0.915424, 0.927621,
   0.951541, 0.95251, 0.963639, 0.971495, 0.97796, 0.98962, 0.992988, 0.997195, 0.998598}}
```

```
xy = Transpose[{NormalCdf[y, Inv], x}]
```

```
{{-1.78534, 238.}, {-1.44516, 244.}, {-0.927042, 268.},
 {-0.890815, 270.}, {-0.305415, 286.}, {-0.149709, 291.}, {0.180083, 302.},
 {0.487687, 316.}, {0.598033, 328.}, {0.608776, 329.}, {0.777512, 351.},
 {0.839952, 376.}, {1.06145, 421.}, {1.06709, 422.}, {1.37493, 454.},
 {1.4583, 470.}, {1.65998, 501.}, {1.66969, 507.}, {1.79457, 563.}, {1.90324, 594.},
 {2.01334, 617.}, {2.31232, 670.}, {2.45666, 719.}, {2.76978, 784.}, {2.98837, 846.}}
```

```
PlotData[xy];
```
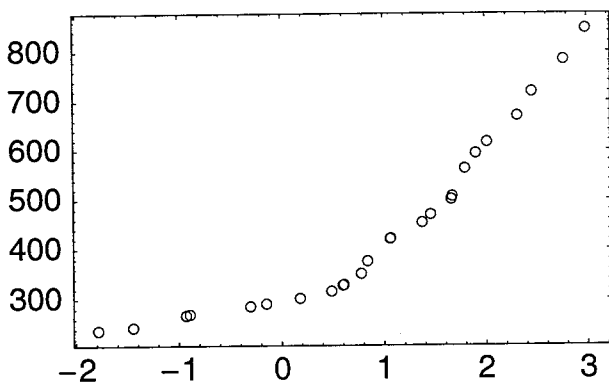

end

▪ sabwTWA

```
{nsam, nsim} = {2000, 10};
Clear[fxn];
fxn[a1_, a2_, a3_] := Plus@@ ({a1, a2, a3} {
      2, 16, 52}/70)
```

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim];]
```

{233.983 Second, Null}

```
TBL/@jen
```

Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	Pval
-0.000142045	0.017043	0.775025	3	0.855431

Fractile	Value	CVM(%)
0.01	259.77	0.362611
0.05	279.448	0.0700455
0.5	326.046	0.0251131
0.95	373.11	0.0744586
0.99	393.457	0.15432
Mean	325.884	$0. + 4.93556 \times 10^{-7}$ I
Variance	777.661	0.230742

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → 225, Xmax → 430];
```
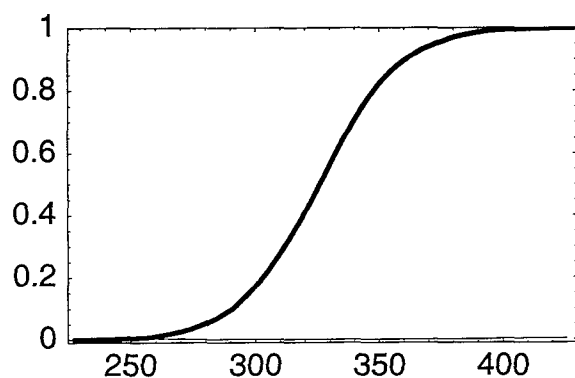


```
Sqrt[777.660703210145509]
```

```
27.8865685090537002
```

```
{EV[cdf, Empirical -> True], SD[cdf, Empirical -> True], Idf[cdf, {.5, .95}]}
```

```
{325.884, 27.8746, {326.046, 373.11}}
```

```
sdf = StandardizeCdf[cdf, 404];
```

```
{EV[sdf, Empirical -> True], Idf[sdf, {.5, .95}]}
```

```
{325.875, {326.046, 373.235}}
```

```
WriteMatrix["Bogen's:Desktop Folder:sabwratioTWA.txt", sdf];
```

```
{x, y} = Transpose[Rest[Drop[cdf, -1]]];
```

```
xy = Transpose[{NormalCdf[y, Inv], x}];
fit = FIT[xy, {1, X}, X, Report → True];
```

```
Coef       LS Est.      SD           95%LCL       95%UCL
q[0]       325.884      0.0240524    325.837      325.931
q[1]       27.9467      0.0241325    27.8994      27.994
```

$R^2 = 0.998512$

$F(1,1998) = 1.34108 \times 10^6$ 2-tail p= $3.7487211392 \times 10^{-2827}$

```
PlotData[xy, FitTo → {fit . {1, X}, X}];
```
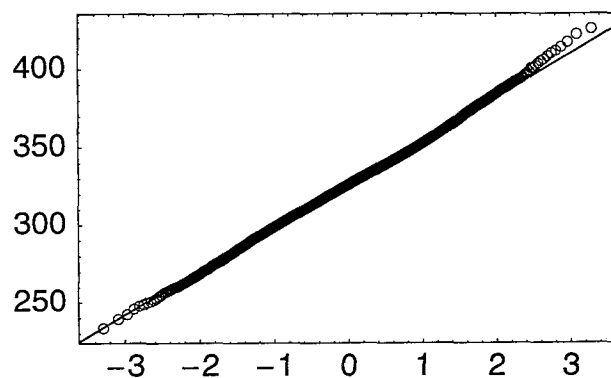
end

end

■ Inhalation (L/kg-d) OHEA. 1996. Stochastic Analysis, p. 3-31 - 3-32.

```
pval = .01 {1, 5, 10, 25, 50, 75, 90, 95, 99, 100 (1 - 1579^-1)};
zval = NormalCdf[pval, Inverse];
yAdult = {112.8, 171.4, 179.7, 185.2, 206, 245.6, 295.1, 366.6, 494, 638.8};
yChild = {342.5, 364.5, 375, 401.5, 441, 489.5, 540.5, 580.5, 663.3, 747.5};
{zyA, zyC} = Transpose[{zval, Log[#]}]&/@{yAdult, yChild};
```

Option 1 (Not Used): Calculate lognormal parameters from OHEA data

```
Clear[x];
{my, sy} = FIT[zyA, {1, x}, x];
{gm, gsd} = E^{my, sy};
{{gm, gsd}, MSDx[gm, gsd]}
```

```
{{231.574, 1.32974}, {241.172, 70.1495}}
```

```
PlotData[zyA, FitTo -> {my + sy * x, x}];
```
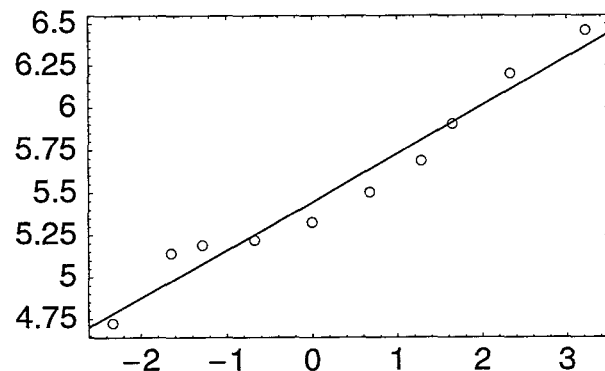


```
{my, sy} = FIT[zyC, {1, x}, x];
{gm, gsd} = E^{my, sy};
{{gm, gsd}, MSDx[gm, gsd]}
```

```
{{456.675, 1.15697}, {461.555, 67.6548}}
```

```
PlotData[zyC, FitTo -> {my + sy * x, x}];
```
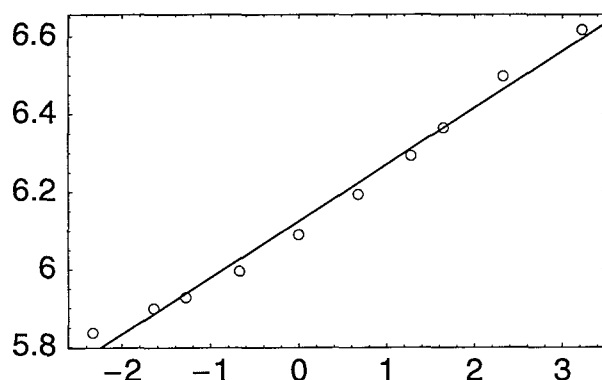

Option2 (Used): Use lognormal parameters derived from reported OHEA mean and SD values for Adult and Child distributions

```
{mA, sdA, mC, sdC} = {225.2, 64.634, 452, 67.73};
{{gmA, gsdA} = GMGSDx[225.2, 64.634],
{gmC, gsdC} = GMGSDx[452, 67.73],
{cvA, cvC} = {sdA / mA, sdC / mC}}

{{216.461, 1.32491}, {447.009, 1.16069}, {0.287007, 0.149845}}

{inhA, inhC} = SimulateCdf[{{{LN, Log[{gmA, gsdA}]}, {LN, Log[{gmC, gsdC}]}}, 5000];
cdfs = Cdf /@ {inhA, inhC};
PlotCdf[Reverse[cdfs]];

Output-Sample Rank-Correlation Matrix:

1.          0.000314
0.000314  1.

Jennrich's Asymptotic Chi-Square Test of Homogeneity

   Between Input & Target Correlation Matrices

   For 5000 2-Variate Normal Samples:

Chi2(1)= 0.000492816    1-tail p= 0.982289  (NS)
```
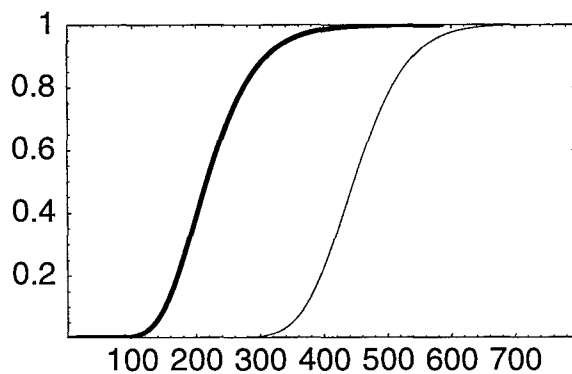

From Finley et al. 1994 (CalEPA 1996, p. 10-7), the BW distribution for adult males & females is ~LN and CV[BW] = ~0.22:

```
{ev, sd} = {71., 15.9};
{cv = sd / ev, {BWgm, BWgsd} = GMGSDx[ev, sd], Log[{BWgm, BWgsd}]}
```

```
{0.223944, {69.2839, 1.24759}, {4.23821, 0.22121}}
```

From CalEPA/OHEA (1996, Stochastic Analysis, p. 3-31 - 3-32; *cit.* above), cvA = CV[24*Q/BW] = CV[Qtot/BW] = ~ 0.3, where Q denotes total ventilation rate in L/h. From Allen and Fisher (1993), alveolar ventilation rate in L/h is modeled as Q ~ 12.9*BW^0.7, and Qtot ~ kQ for some constant k. Now let VQ be LN-distributed with an arithmetic mean of 1, where VQ represents variation in Q not attributable to that in BW. Thus, Q ~ 12.9*VQa*BW^.7, whence Q/BW ~ 12.9*VQ*BW^-0.3. It follows from the method of moments that CV[BW^-.3] = CV[BW^.3] = 0.06644, whence CV[VQ] = 0.2919, GM[VQ] = 0.9599, GSD[VQ] = 1.331, Log[{GM[VQ], GSD[VQ]}] = {-0.0408868, 0.285961}.

```
{{o1, o2} = MSDx[a BWgm^-.3, BWgsd^-.3], cvBW3 = o2 / o1}
```

```
{{0.281039 a, 0.0186711 a}, 0.066436}
```

```
{o = GMGSDx1[0.3, cvBW3], {muyX, sdyX} = Log[o]}
```

```
{{0.959938, 1.33104}, {-0.0408868, 0.285961}}
```

```
{{mX, sdX} = MSDx@@o, cvX = sdX / mX}
```

```
{{1., 0.291908}, 0.291908}
```

```
(* By definition, the CV of (BW^.3 * X) = *)
{Sqrt[E^ (sdyX^2 + Log[BWgsd^.3]^2) - 1],
((1 + cvBW3^2) (1 + cvX^2) - 1)^.5}
```

```
{0.3, 0.3}
```

■ InhaleALL

```
{f12, f12p}
```

```
{0.172193, 0.827807}
```

```
adf = AverageCdf[cdfs, Weights -> {f12p, f12}];
sadf = N[StandardizeCdf[adf, 404]];
```

```
PlotCdf[sadf, Xmin -> -.001, Ymin -> -.001];
```
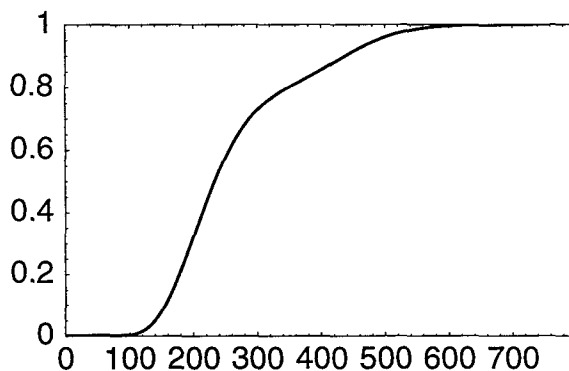
```
{EV[#], SD[#], Idf[#, {.5, .95}]}&/@{adf, sadf}
```

{{264.165, 107.491, {233.106, 487.138}}, {264.165, 108.051, {233.106, 487.146}}}

```
N[{{pA, pC} = Edf[#, 487.138]&/@cdfs, wp = {f12p*pA, f12*pC}, Plus@@wp}, 10]
```

{{0.9982290631, 0.7181436766}, {0.8263408411, 0.1236594327}, 0.9500002739}

$$\left(\text{f12p NormalCdf}\left[\frac{\text{Log}[x/\text{gmA}]}{\text{Log}[\text{gsdA}]}\right] + \text{f12 NormalCdf}\left[\frac{\text{Log}[x/\text{gmC}]}{\text{Log}[\text{gsdC}]}\right]\right) /. \; x \rightarrow \text{Range}[487.3, 487.4, .01]$$

{0.949947, 0.949956, 0.949964, 0.949972, 0.949981, 0.949989, 0.949997, 0.950006, 0.950014, 0.950023}

$$N\left[\text{FindRoot}\left[\text{f12p CDF}\left[\text{NormalDistribution}[0, 1], \frac{\text{Log}[x/\text{gmA}]}{\text{Log}[\text{gsdA}]}\right] + \right.\right.$$
$$\left.\left. \text{f12 CDF}\left[\text{NormalDistribution}[0, 1], \frac{\text{Log}[x/\text{gmC}]}{\text{Log}[\text{gsdC}]}\right] == 95/100, \{x, 485, 480, 490\}\right],\right.$$
$$16\Big]$$

{x → 487.3630111243049}

```
WriteMatrix["BogenHD:Desktop Folder:inhaleALL.txt", sadf];
```

■ InhaleTWA

```
{{f12p, f12}, {f12p, f12} 70}
```

{{0.827807, 0.172193}, {57.9465, 12.0535}}

```
{nsam, nsim} = {2000, 10};
Clear[fxn];
fxn[a1_, a2_] := Plus@@ ({a1, a2} {
     58, 12}/70)
```

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim];]
```

{143.533 Second, Null}

```
TBL/@jen
```

Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	Pval
0.0010747	0.00945733	0.178882	1	0.672335

Fractile	Value	CVM(%)
0.01	166.336	0.275487
0.05	187.965	0.14204
0.5	257.154	0.0674972
0.95	362.928	0.17087
0.99	424.294	0.272136
Mean	263.865	7.46556×10^{-7}
Variance	2937.4	0.0463849

```
PlotCdf[{cvm, cdf}, Ymin -> -.01, Xmin -> 95, Xmax -> 430];
```
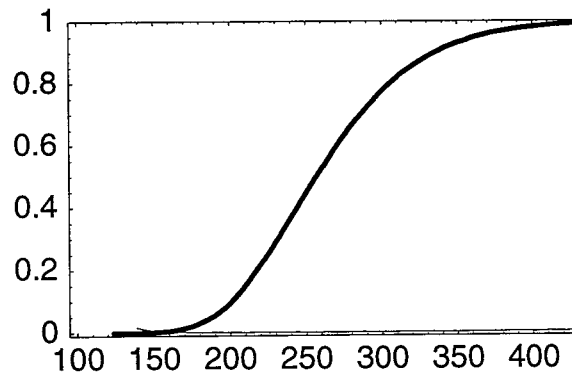


```
{EV[cdf, Empirical -> True], SD[cdf, Empirical -> True], Idf[cdf, {.5, .95}]}

{263.865, 54.1887, {257.154, 362.928}}


sdf = StandardizeCdf[cdf, 404];
{EV[sdf, Empirical -> True], SD[sdf, Empirical -> True], Idf[sdf, {.5, .95}]}

{263.752, 53.5784, {257.154, 362.847}}


WriteMatrix["Bogen's:Desktop Folder:inhaleTWA.txt", sdf];
```

Appendix 3.C

Fraction of Lifetime at One Local Residence

```
<< RiskQ`;

HardDrive= "Bogen";
PathName[filename_, hardDrive_String:HardDrive] := Module[{file= filename},
        If[Head[file] =!=String, file= ToString[file]];
        StringJoin[hardDrive, ":Ken:TCE Air Force:Data:", file]
    ];
```

■ **Israeli, M., and C. Nelson. 1992. Distribution and expected time of residence for U.S. housholds.** *Risk Anal.* **12, 65-72.**

```
St = E^-(a1*b1 (1-E^(-t/b1))+a2 t+a3*b3 (E^(t/b3)-1)) ;
pt = St ((a1* (E^-(t/b1))) +a2+ (a3* (E^ (t/b3))));
Rt = ----'pt----;
     a1+a2+a3
```

Coefficients a and b all have units of y^{-1} and y, respectively, from Israeli and Nelson Table II (All households, W-Rgn)

■ **All households**

```
ruleA= Rule@@ # & /@Transpose[{{a1, b1, a2, a3, b3}, {.1503, 1.88, .0679, .0015, 13.3}}] ;

{St, Rt, RtrA= (Rt /. ruleA)}
```

$$\left\{E^{-a1\,b1\left(1-E^{-\frac{t}{b1}}\right)-a3\,b3\left(-1+E^{\frac{t}{b3}}\right)-a2\,t}, \; \frac{E^{-a1\,b1\left(1-E^{-\frac{t}{b1}}\right)-a3\,b3\left(-1+E^{\frac{t}{b3}}\right)-a2\,t}\left(a2+a1\,E^{-\frac{t}{b1}}+a3\,E^{\frac{t}{b3}}\right)}{a1+a2+a3}, \right.$$

$$\left. 4.55166\,E^{-0.282564\,(1-E^{-0.531915\,t})-0.01995\,(-1+E^{0.075188\,t})-0.0679\,t}\,(0.0679+0.1503\,E^{-0.531915\,t}+0.0015\,E^{0.075188\,t})\right\}$$

```
({St, pt, Rt} /. Append[ruleA, t -> #]) & /@ {0, 50}

{{1, 0.2197, 1.}, {0.0109554, 0.00144923, 0.00659639}}
```

```
Plot[Log[10, Evaluate[St /. ruleA]], {t, 0, 50},
AxesOrigin -> {0, -3.01}, PlotRange -> {{0, 50}, {-3.01, 0}}, Frame -> True];
```
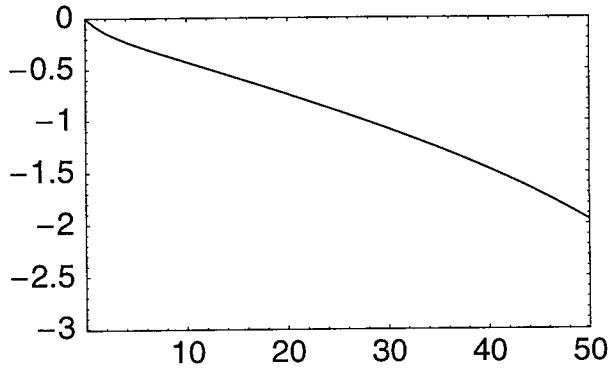


```
Plot[Log[10, Evaluate[RtrA]], {t, 0, 50},
AxesOrigin -> {-0.01, -3.01}, PlotRange -> {{-.01, 50}, {-3.01, 0}}, Frame -> True];
```
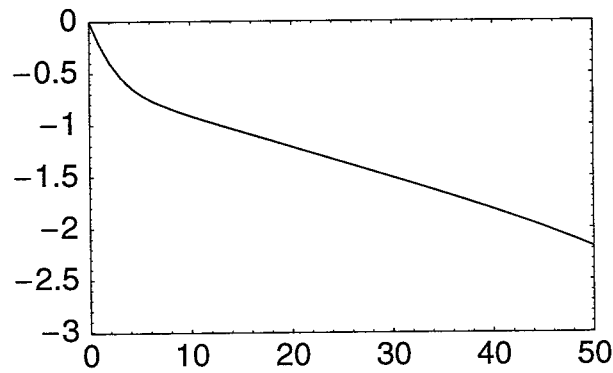

■ Western Region

```
ruleW = Rule@@#& /@Transpose[{{a1, b1, a2, a3, b3}, {.2029, 1.74, .0832, .0008, 10.3}}]
```

$\{a1 \to 0.2029, b1 \to 1.74, a2 \to 0.0832, a3 \to 0.0008, b3 \to 10.3\}$

```
RtrW = (Rt /. ruleW)
```

$3.48554\, E^{-0.353046\,(1-E^{-0.574713\,t})-0.00824\,(-1+E^{0.0970874\,t})-0.0832\,t}\,(0.0832 + 0.2029\,E^{-0.574713\,t} + 0.0008\,E^{0.0970874\,t})$

```
({St, pt, Rt} /. Append[ruleW, t -> #])& /@ {0, 50}
```

$\{\{1, 0.2869, 1.\}, \{0.0038411, 0.00071383, 0.00248808\}\}$

```
Plot[Evaluate[Log[10, (St /. #)&/@{ruleA, ruleW}]], {t, 0, 50},
AxesOrigin -> {0, -3.01}, PlotRange -> {{0, 50}, {-3.01, 0}}, Frame -> True];
```
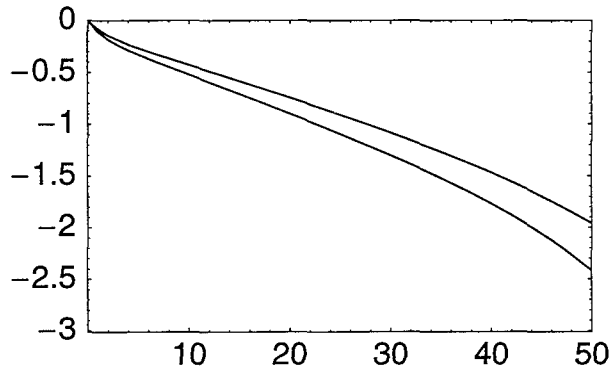


```
Plot[Evaluate[Log[10, {RtrA, RtrW}]], {t, 0, 50},
AxesOrigin -> {-0.01, -3.01}, PlotRange -> {{-.01, 50}, {-3.01, 0}}, Frame -> True];
```
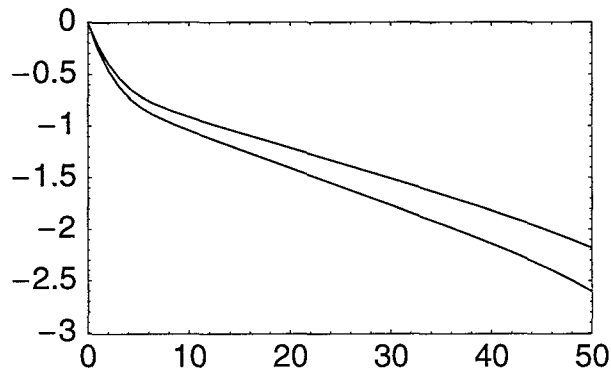

■ Adaptation of model to account for fraction Fm of moves that are out of a Western-region water distribution system

```
{RtrW, (RtrW /. t -> #)&/@{0, 50, 70}}
```

$$\left\{3.48554\ E^{-0.353046\ (1-E^{-0.574713\ t})-0.00824\ (-1+E^{0.0970874\ t})-0.0832\ t}\ (0.0832 + 0.2029\ E^{-0.574713\ t} + 0.0008\ E^{0.0970874\ t}),\right.$$

$$\left\{1., 0.00248808, 3.67283 \times 10^{-6}\}\right\}$$

$$\left\{\text{fmLo} = \frac{1}{3} + \sqrt{\frac{2}{3}\frac{1}{3}0.05}\ ,\ \text{fmHat} = \frac{2}{3}\right\}$$

$$\left\{0.438743,\ \frac{2}{3}\right\}$$

```
time = Join[Range[0, 10, .1], Range[11, 50], Range[55, 70, 5]];
Mt = ((1 - RtrW¹) /. t -> time);
MtFmHat = ((1 - RtrW^fmHat) /. t -> time);
MtFmLo = ((1 - RtrW^fmLo) /. t -> time);

{cdfTR, cdfTRFmHat, cdfTRFmLo} =
 Append[Drop[Transpose[{time, #}], -1], {70, 1}]&/@{Mt, MtFmHat, MtFmLo};
```

```
RtrW /. t -> 1
```

```
0.544555
```

```
PlotData[{cdfTRFmLo, cdfTRFmHat, cdfTR}, Xmin -> -.01, Xmax -> 70, Ymin -> 0, Ymax -> 1,
 DotSize -> {.0001, .008, .0125}, Style -> 0,
 JoinPoints -> True, FitTo -> {RtrW, t}];
```
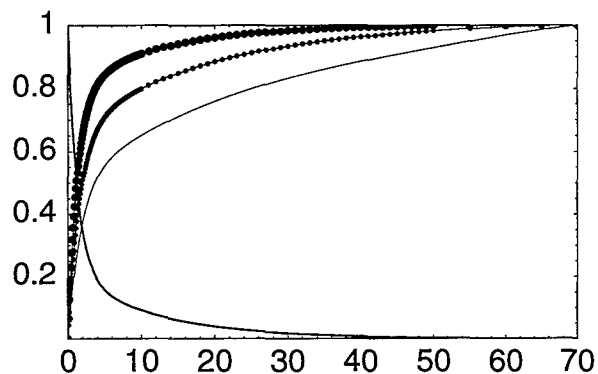


```
{EV[cdfTR], Sqrt[Var[cdfTR]], Idf[cdfTR, {.025, .5, .95, .975}]}
```

```
{3.48741, 6.83815, {0.0378193, 1.16603, 17.0596, 25.4481}}
```

```
{EV[cdfTRFmHat], Sqrt[Var[cdfTRFmHat]], Idf[cdfTRFmHat, {.025, .5, .95, .975}]}
```

```
{7.02862, 11.8589, {0.0560921, 1.9295, 35.0626, 46.1341}}
```

```
{EV[cdfTRFmLo], Sqrt[Var[cdfTRFmLo]], Idf[cdfTRFmLo, {.025, .5, .95, .975}]}
```

```
{12.9946, 17.7241, {0.0845741, 3.63964, 55.2843, 61.7931}}
```

```
Rt = (RtrW /. t -> time);
xy = Append[Drop[Transpose[{time, Rt}], -1], {70, 0}];
IRt = Interpolation[xy, InterpolationOrder -> 1];
```

```
10.5 / 16. (* = US fraction moving to same county *)
```

```
0.65625
```

```
fms = SimulateCdf[{Tri, {1 / 3, 2 / 3, 1}}, 2000];
Idf[Cdf[fms], .05]
```

```
0.438716
```

```
PlotCdf[Cdf[fms]];
```
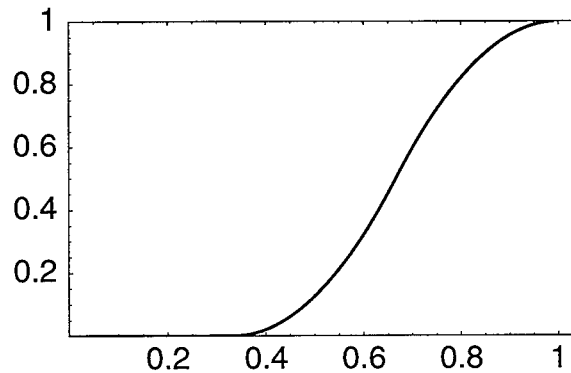


```
tRange = Prepend[time, t];

Tbar = NIntegrate[IRt[t]^#, Evaluate[tRange]]&/@fms;

cdfTbar = Cdf[Tbar];

PlotCdf[cdfTbar, Xmin -> -.01, Xmax -> 16];
```
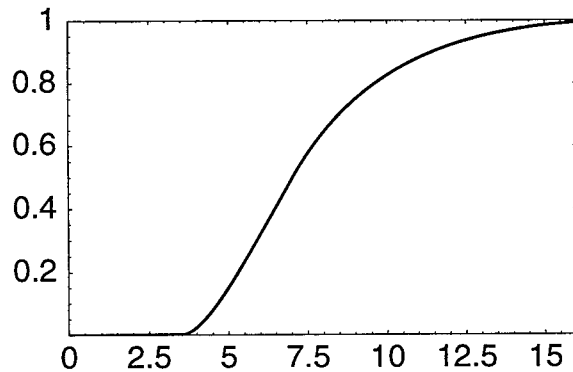


```
{EV[Tbar], Sqrt[Var[Tbar]], Idf[cdfTbar, {.025, .5, .975}]}

{7.5984, 2.67068, {4.00903, 7.03101, 14.2792}}

D[a^Fm, {Fm, 2}]

a^Fm Log[a]^2

cdfFm = TriangularCdf[1 / 3, 2 / 3, 1, 500];
{EV[cdfFm], Var[cdfFm]}

{0.666667, 0.0185294}

cdfFm = TriangularCdf[1 / 3, 2 / 3, 1, 10000];
{EV[cdfFm], Var[cdfFm]}

{0.666667, 0.0185186}
```

```
0.0185186464965580999^-1
```

```
53.999626818615568
```

```
Rti = Transpose[xy][[2]];
fmBar = 2 / 3;
fmVar = 1 / 54;
```

$$pi = 1 - \left((Rti \wedge fmBar) \left(1 + \frac{Log[Rti]^2}{2} \ fmVar \right) \right);$$

```
pi = Append[Drop[pi, -1], 1];
```

```
cdfTangbr = Transpose[{time, pi}];
```

```
PlotCdf[cdfTangbr, Xmin -> -.01, Xmax -> 70];
```
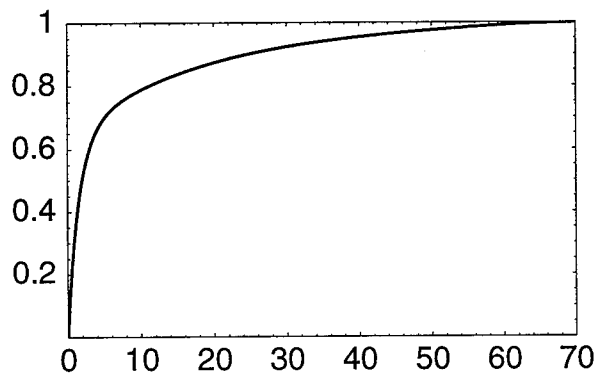


```
PlotCdf[{cdfTangbr, cdfTbar}, Xmin -> -.01, Xmax -> 70];
```
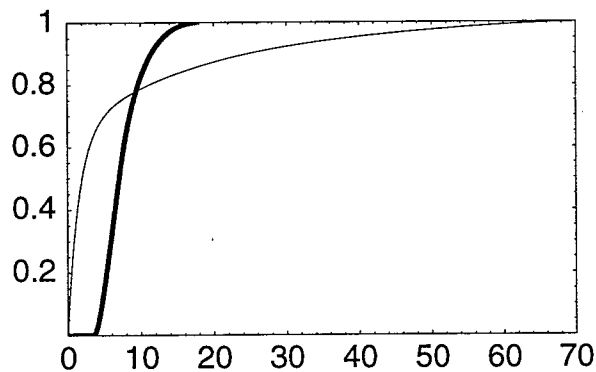


```
{EV[cdfTangbr], Sqrt[Var[cdfTangbr]], Idf[cdfTangbr, {.025, .5, .975}]}
```

```
{7.55321, 12.748, {0.0561443, 1.96772, 49.7245}}
```

```
sdfTbar = StandardizeCdf[cdfTbar, time];
```

```
TBL[out = Prepend[Transpose[{time, N[Last/@sdfTbar, 8],
    N[Last/@cdfTangbr, 8]}], {Time, FTbar, FTangbr}]]
```

```
Time  FTbar               FTangbr
0     0                   0.
0.1   0.000014065155      0.044528152
0.2   0.00002813031       0.085900584
0.3   0.000042195465      0.12439054
```

0.4	0.00005626062	0.16024284
0.5	0.000070325776	0.19367728
0.6	0.000084390931	0.22489166
0.7	0.000098456086	0.25406435
0.8	0.00011252124	0.2813565
0.9	0.0001265864	0.30691402
1.	0.00014065155	0.33086927
1.1	0.00015471671	0.35334255
1.2	0.00016878186	0.37444339
1.3	0.00018284702	0.39427172
1.4	0.00019691217	0.41291884
1.5	0.00021097733	0.43046832
1.6	0.00022504248	0.4469968
1.7	0.00023910764	0.46257462
1.8	0.00025317279	0.47726652
1.9	0.00026723795	0.49113209
2.	0.0002813031	0.50422633
2.1	0.00029536826	0.51660002
2.2	0.00030943341	0.52830013
2.3	0.00032349857	0.53937017
2.4	0.00033756372	0.54985047
2.5	0.00035162888	0.55977846
2.6	0.00036569403	0.56918892
2.7	0.00037975919	0.57811422
2.8	0.00039382434	0.58658448
2.9	0.0004078895	0.59462778
3.	0.00042195465	0.60227029
3.1	0.00043601981	0.60953645
3.2	0.00045008496	0.61644911
3.3	0.00046415012	0.62302958
3.4	0.00047821527	0.62929784
3.5	0.00049228043	0.63527256
3.6	0.0013753219	0.64097123
3.7	0.0046798441	0.64641023
3.8	0.0097212226	0.65160492
3.9	0.016294389	0.6565697
4.	0.024222899	0.66131805
4.1	0.033351911	0.66586267
4.2	0.043545234	0.67021542
4.3	0.054682977	0.67438749
4.4	0.066658284	0.67838933
4.5	0.079376907	0.68223081
4.6	0.092754949	0.68592116
4.7	0.10671755	0.68946908
4.8	0.12119799	0.69288271
4.9	0.13613655	0.69616975
5.	0.15147976	0.69933738
5.1	0.16717972	0.70239239
5.2	0.18319323	0.70534116
5.3	0.19948154	0.70818966
5.4	0.21600975	0.71094353
5.5	0.23274635	0.71360807
5.6	0.24966287	0.71618824
5.7	0.26673362	0.71868873
5.8	0.28393535	0.72111394
5.9	0.30124702	0.72346801
6.	0.31864953	0.72575483
6.1	0.33612562	0.72797806
6.2	0.3536596	0.73014114
6.3	0.37123722	0.73224731
6.4	0.38884558	0.73429961
6.5	0.40647292	0.7363009
6.6	0.4241086	0.73825388
6.7	0.44174292	0.74016109
6.8	0.45936707	0.74202489
6.9	0.47697308	0.74384753
7.	0.49455365	0.74563113
7.1	0.51196343	0.74737765
7.2	0.52877535	0.74908897
7.3	0.54498488	0.75076684
7.4	0.56061815	0.75241291
7.5	0.57570017	0.75402873
7.6	0.59025445	0.75561577
7.7	0.60430319	0.75717539
7.8	0.61786745	0.7587089
7.9	0.63096716	0.7602175
8.	0.6436211	0.76170234
8.1	0.65584727	0.76316451
8.2	0.66766265	0.76460501
8.3	0.67908339	0.7660248
8.4	0.69012475	0.76742478

```
8.5    0.70080134    0.76880579
8.6    0.71112707    0.77016864
8.7    0.72111508    0.77151407
8.8    0.73077791    0.7728428
8.9    0.74012765    0.77415548
9.     0.74917561    0.77545275
9.1    0.75793282    0.77673519
9.2    0.7664095     0.77800336
9.3    0.77461566    0.7792578
9.4    0.78256085    0.78049899
9.5    0.79025397    0.7817274
9.6    0.79770389    0.78294347
9.7    0.80491887    0.78414763
9.8    0.81190674    0.78534025
9.9    0.81867513    0.78652171
10.    0.82523141    0.78769236
11     0.88053089    0.79886871
12     0.9207822     0.80923462
13     0.9498595     0.81893815
14     0.97047554    0.82807078
15     0.98457918    0.83669313
16     0.99360366    0.84484873
17     0.99861864    0.8525714
18     1.            0.85988918
19     1.            0.86682654
20     1.            0.87340553
21     1.            0.8796464
22     1.            0.88556805
23     1.            0.89118824
24     1.            0.89652371
25     1.            0.90159033
26     1.            0.90640318
27     1.            0.91097659
28     1.            0.91532422
29     1.            0.91945913
30     1.            0.92339381
31     1.            0.9271402
32     1.            0.9307098
33     1.            0.93411365
34     1.            0.93736239
35     1.            0.94046627
36     1.            0.9434352
37     1.            0.94627875
38     1.            0.94900617
39     1.            0.9516264
40     1.            0.95414801
41     1.            0.95657927
42     1.            0.95892806
43     1.            0.96120185
44     1.            0.96340761
45     1.            0.9655518
46     1.            0.96764026
47     1.            0.96967807
48     1.            0.97166952
49     1.            0.97361794
50     1.            0.97552559
55     1.            0.98444368
60     1.            0.99193758
65     1.            0.99706939
70     1.            1.
```

Put[out, PathName[Tbarang]];

Appendix 3.D
Effective Genotoxic Dose

```
<< RiskQ`;
HardDrive = "Bogen";
PathName[filename_, hardDrive_String: HardDrive] := Module[{file = filename},
        If[Head[file] =!= String, file = ToString[file]];
        StringJoin[hardDrive, ":Ken:TCE Air Force:Data:", file]];
```

■ Input Empirical (Derived) Distributions. Recreate Input Distributions from Phase I study. Exposure in mg/kg-d.

■ Log-Transform Utility Functions

```
GMGSDx::usage =
  "GMGSDx[Mx,SDx] returns the  geometric mean and geometric sdandard deviation of
    a lognormal variate X that also has the specified arithmetic mean Mx and
    arithmetic sdandard deviation SDx, based on the method of moments.";

MSDx::usage =
  "MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation
    of a lognormal variate X that also has the specified geometric mean GMx and
    geometric sdandard deviation GSDx, based on the method of moments.";

GMGSDx1::usage =
  "GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that
    the product X1*X2 has the desired coefficient of variation (CV) = cvWant,
    conditional on the lognormal variate X2 having an arithmetic mean and
    CV equal to 1 and cv2, respectively, based on the method of moments.";

MSDx[GMx_, GSDx_] := Module[{mux, sigy},
        sigy = Log[GSDx];
        mux = GMx E ^ ((sigy^2) / 2);
        mux {1, Sqrt[E ^ (sigy^2) - 1]}]]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
        sigy = Sqrt[Log[1 + (SDx / Mx) ^ 2]];
        muy = Log[Mx] - (sigy^2) / 2;
        E ^ {muy, sigy}]
```

$$
\text{GMGSDx1}[cvWant\_, cv2\_] := \text{Module}\Big[\{my1\},
$$
$$
muy1 = \text{Log}\Big[\text{Sqrt}\Big[\frac{cv2^2 + 1}{cvWant^2 + 1}\Big]\Big];
$$
$$
E ^ \{muy1, \text{Sqrt}[-2\,muy1]\}\}
$$
$$
\Big] \; /; \; cvWant >= cv2
$$

```
SABW = ToExpression[ReadList[PathName["sabwratioALL.txt"], Word, RecordLists -> True]];
Inhale = ToExpression[ReadList[PathName["inhaleALL.txt"], Word, RecordLists -> True]];
    (*Note: Inhale in L/kg-d *)
Conc = ToExpression[ReadList[PathName["concentration.txt"], Word, RecordLists -> True]];
    (*Note: Conc in mg/L *)
Tbarang = Rest[Get[PathName[Tbarang]]];
TresBar = #[[{1, 2}]]&/@Tbarang;
TresAng = #[[{1, 3}]]&/@Tbarang;
Dimensions/@{SABW, Inhale, Conc, TresBar, TresAng}
```

```
{{27, 2}, {405, 2}, {405, 2}, {145, 2}, {145, 2}}
```

```
EV/@{SABW, Inhale, Conc, TresBar, TresAng}
```

```
{325.881, 264.032, 0.0229323, 7.59939, 7.55321}
```

```
{BW, Vmet, VQ} =
 {{LN, {4.23821, 0.22121}},
  {LN, {-0.15154, 0.550528}}, {LN, {-0.0408868, 0.285961}}};
```

■ Constants

```
{TresBarAng, TresAngBar} = EV[#, Empirical -> True]&/@{TresBar, TresAng}
```

```
{7.59939, 7.55321}
```

$$\text{inhalebar} = 12.9 * 71^{0.74-1} * \frac{24}{1000} \qquad (* \text{ m3/kg-d } *)$$

```
0.102205
```

```
EFcon  =       350;    (* d/y  *)
ATcon  =       25550;  (* d    *)
ConcAng =      0.0223; (* mg/L *)
IngestBar=     0.0242; (* L/kg-d *)
InhaleBar =    0.102;  (* m3/kg-d *)
SABWBar  =     325.881;(* cm2/kg *)
TresBarAng=    7.59358;(* y *)
TresAngBar=    7.55321;(* y *)
```

end

■ Fractions metabolized (summary–see "E. Effective Cytotoxic Dose")

Correlation between Vmet and Fmo = 0.86
Correlation between VQ and Fmr = -0.75
Correlation between Vmet and Fmr = 0.45

```
Fmo = Get[PathName[Fmo]]; Fmr = Get[PathName[Fmr]];
CdfQ/@{Fmo, Fmr}
```

```
{True, True}
```

```
{FmoBar, FmrBar} = {0.888543, 0.6732836};
```

end

Note:
All distributions below are multiplied by Scale→1000

■ Ingestion

■ EingBar = Uncertainty in Population-Average Level

{TresBar,Conc} = uncertain variates

```
{nsam, nsim} = {2000, 10};
cdfs = {TresBar, Conc};
Clear[fxn];
                         EFcon
fxn[t_, c_] := IngestBar * t ───── c * FmoBar
                         ATcon

fxn[TresBarAng, ConcAng]
```

0.0000498795

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1000];]
```

{200.817 Second, Null}

```
TBL /@ jen
```

					Fractile	Value	CVM(%)
					0.01	0.0143149	2.23118
					0.05	0.0230076	0.210285
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	Pval	0.5	0.0429802	0.168471
-0.00355814	0.0158828	0.504529	1	0.477517	0.95	0.0889072	0.414161
					0.99	0.134017	1.78008
					Mean	0.0507837	0.514653
					Variance	0.00710812	18.3718

```
{First[cdf], Last[cdf]}
```

{{0.000853119, 0}, {4.43688, 1}}

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .125];
```
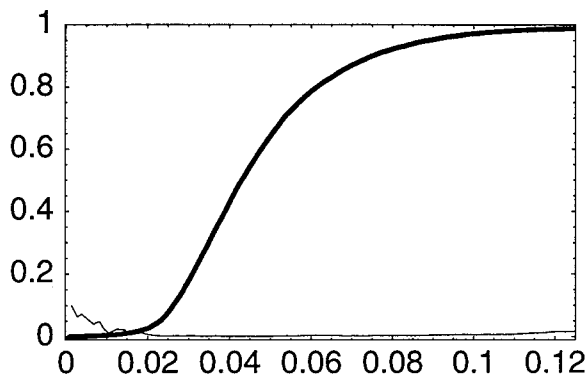


```
Put[cdf, PathName[EingestBar]];
```

end

■ <Eing> = Variability in Expected Level

{Ingest,TresAng,Fmo} = heterogeneous variates

```
{nsam, nsim} = {2000, 10};
Ingest = {LN, {Log[.0198], Log[1.88]}};
cdfs = {Ingest, TresAng, Fmo};
Clear[fxn];
                       EFcon
fxn[ing_, t_, f_] := ing * t ─────── ConcAng * f
                       ATcon

fxn[IngestBar, TresAngBar, FmoBar]

0.0000496144

Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1000];]

{47.8167 Second, Null}

TBL/@jen
```

						Fractile	Value	CVM(%)
						0.01	0.000106569	3.27128
						0.05	0.000509143	1.43336
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj		0.5	0.0108513	0.866917
0.00168143	0.0152108	0.522666	3	0.995174		0.95	0.23658	1.24671
						0.99	0.531411	4.21273
						Mean	0.0491639	0.747214
						Variance	0.0111399	5.47589

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .4];
```
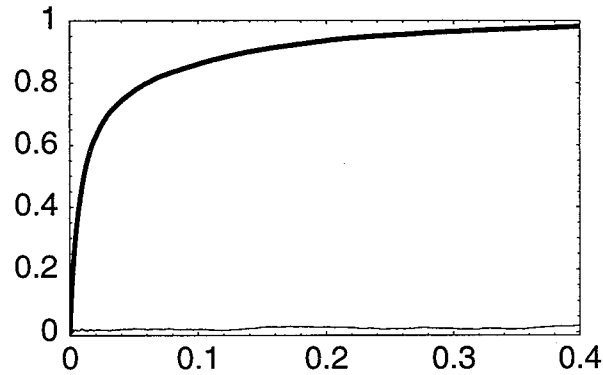


```
Put[cdf, PathName[EingestAng]];
```

end

end

■ Inhalation Exposure

■ EinhBar = Uncertainty in Population-Average Level

{TresBar,Conc} = uncertain variates

```
{AEshHBar, AEbaHBar, AEhHBar} = 1 / (EV /@ {
{1. / #[[1]], #[[2]]}& /@ RQ[Cdf, U, {4, 20}, 2000],
{1. / #[[1]], #[[2]]}& /@ RQ[Cdf, U, {10, 100}, 2000],
{1. / #[[1]], #[[2]]}& /@ RQ[Cdf, U, {300, 1200}, 2000]})
```

{9.94136, 39.0865, 649.213}

```
cdfs = {TresBar, Conc};
```

$$\text{fxn} \to \text{FmrBar}\left(\frac{12.9}{1000} * 71^{.74-1} * 1\right) * t\ \frac{\text{EFcon}}{\text{ATcon}}\ c\ \frac{1}{24}$$

$$\left(480\ (.76)\left(\frac{.129}{\text{AEshHBar}} + \frac{.33}{\text{AEbaHBar}}\right) + 42\left(.76\ \frac{.54}{.7}\right)\frac{14}{\text{AEhHBar}}\right)$$

fxn → 0.0000136563 c t

```
Clear[fxn, jen, cdf, cvm];
fxn[t_, c_] := 0.0000136563095583647187` c t
```

```
fxn[t, c] /. {c → ConcAng, t → TresBarAng}
```

2.31252×10^{-6}

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, 2000, 10, Scale → 1000];]
```

{205.55 Second, Null}

```
TBL/@jen
```

```
⌠ Mean[Δr]      Max[|Δr|]   JennrichChi2  DegFr  Pval
⌡ 0.000642353   0.0113889   0.259415      1      0.610522 ′
```

```
Fractile  Value         CVM(%)
0.01      0.000683362   2.05913
0.05      0.00106135    0.294054
0.5       0.00198916    0.19258  ⌉
0.95      0.00412975    0.389326 ⌋
0.99      0.00593213    1.01573
Mean      0.00235084    0.583042
Variance  0.0000146843  23.1052
```

```
{First[cdf], Last[cdf]}
```

```
{{0.0000407589, 0}, {0.209696, 1}}
```

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .005];
```
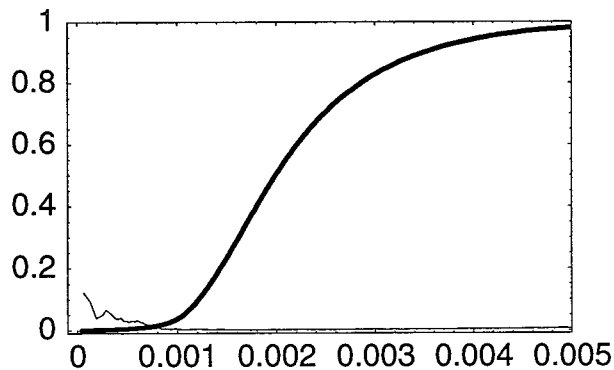


```
Put[cdf, PathName[EinhaleBar]];
```

end

■ <Einh> = Variability in Expected Level

{Inhale,TresAng,Wshower,Whouse,ETshower,ETbath,EThouse,AEshower,AEbath,AEhouse,
T13,T13,VQ,BW} = heterogeneous variates

```
gmgsd = GMGSDx@@#&/@N[{{480, 160}, {42, 15}, {.129, .052}, {.33, .22}}]
```

```
{{455.368, 1.38347}, {39.5532, 1.41408}, {0.119645, 1.47407}, {0.274577, 1.83382}}
```

```
{Wshower, Whouse, ETshower, ETbath} = ({LN, #}&/@Log[gmgsd])
```

```
{{LN, {6.12111, 0.324593}}, {LN, {3.67765, 0.346479}}, {LN, {-2.12323, 0.388026}},
 {LN, {-1.29253, 0.606403}}}
```

```
{AEshower, AEbath, AEhouse, EThouse} =
 ({U, #}&/@{{4, 20}, {10, 100}, {300, 1200}, {8, 20}})
```

```
{{U, {4, 20}}, {U, {10, 100}}, {U, {300, 1200}}, {U, {8, 20}}}
```

```
{InhaleBar, TresAngBar, ConcAng}
```

```
{0.102, 7.55321, 0.0223}
```

```
{VQ, BW}
```

```
{{LN, {-0.0408868, 0.285961}}, {LN, {4.23821, 0.22121}}}
```

Correlation between VQ and Fmr = -0.75 (see "E. Effective Cytotoxic Dose")

```
corr = Table[Table[0, {j}], {j, 13}];
corr = ReplacePart[Reverse[corr], -.75, {1, 1}]
```

```
{{-0.75, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
 {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0},
 {0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0},
 {0, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}}}
```

```
Clear[fxn];
T13 = {T, 13};
cdfs = {Fmr, VQ, BW, TresAng, Wshower,
   Whouse, ETshower, ETbath, EThouse, AEshower, AEbath, AEhouse, T13, T13};
fxn[f_, vq_, bw_, t_, wsh_, wh_, etsh_, etba_, eth_, aesh_, aeba_, aeh_, t13sh_, t13h_] :=
```

$$f * \left(\frac{12.9}{1000} * bw^{.74-1} * vq\right) t \frac{EFcon}{ATcon} ConcAng \frac{1}{24} \left(\right.$$
$$\left. wsh \, (.76 + .029 \, t13sh) \left(\frac{etsh}{aesh} + \frac{etba}{aeba}\right) + wh \left((.76 + .029 \, t13h) \frac{.54}{.7}\right) \frac{eth}{aeh}\right)$$

```
fxn[FmrBar, 1, 71, TresAngBar, 480, 42, .129, .33, 14, AEshHBar, AEbaHBar, AEhHBar, 0, 0]
```

$$2.30022 \times 10^{-6}$$

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, 2000, 10, Correlate → corr, Scale → 1000];]
```

```
{277.6 Second, Null}
```

```
TBL/@jen
```

Mean[Δr]	Max[$|\Delta r|$]	JennrichChi2	DegFr	P-adj
-0.000310836	0.0321802	10.4729	91	1.

Fractile	Value	CVM(%)
0.01	4.8772×10^{-6}	2.95986
0.05	0.0000238177	1.50669
0.5	0.000489184	0.794377
0.95	0.010794	1.05828
0.99	0.0233732	2.245
Mean	0.00225019	0.501055
Variance	0.0000244482	5.8905

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .0125];
```
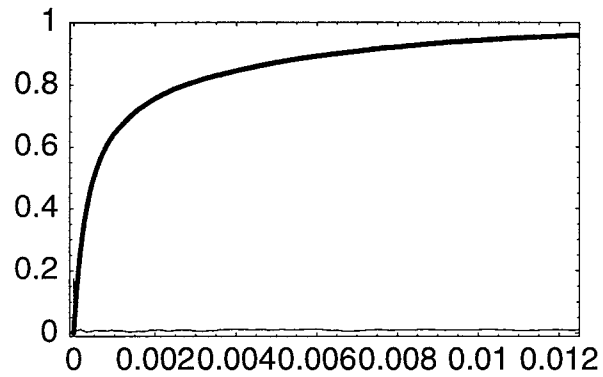


```
Put[cdf, PathName[EinhaleAng]];
```

end

end

■ Dermal Exposure

■ EdermalBar = Uncertainty in Population-Average Level

{TresBar,Conc} = uncertain variates

```
ETshower
```

```
{LN, {-2.12323, 0.388026}}
```

```
cdfs = {TresBar, Conc};
```

$$\text{fxn} \to \text{FmrBar} * \text{SABWBar} * .65 * .263 * .129 * t \, \frac{\text{EFcon}}{\text{ATcon}} \, c \left(1 - \frac{.76}{2}\right) 10^{-3}$$

```
fxn → 0.0000410946 c t
```

```
Clear[fxn];
fxn[t_, c_] := 0.0000410945965500063259` c t
```

```
fxn[t, c] /. {c → ConcAng, t → TresBarAng}
```

$$6.95883 \times 10^{-6}$$

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, 2000, 10, Scale → 1000];]
```

```
{33.3833 Second, Null}
```

```
TBL/@jen
```

```
{ Mean[Δr]     Max[|Δr|]   JennrichChi2  DegFr   P-adj
{-0.00280571  0.0144425   0.41717        1       0.929433 '
```

```
Fractile  Value       CVM(%)
0.01      0.00213724  1.34291
0.05      0.00318948  0.37956
0.5       0.005998    0.123699
0.95      0.0123634   0.459695
0.99      0.0187893   1.62332
Mean      0.0071258   0.269769
Variance  0.000136076 8.27214
```

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .015];
```
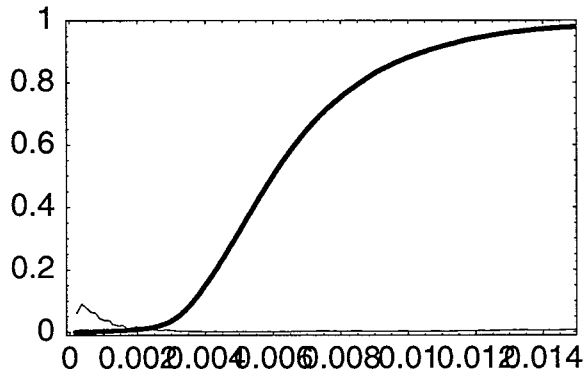


```
Put[cdf, PathName[EdermalBar]];
```

end

■ \<Edermal\> = Variability in Expected Level

{SABW,Fs,Kp,ETshower,TresAng,T13,Fmr} = heterogeneous variates

```
CdfQ/@{SABW, TresAng}
```

```
{True, True}
```

```
{ConcAng, Fs, Kp, ETshower, T13}
```

```
{0.0223, Fs, Kp, {LN, {-2.12323, 0.388026}}, {T, 13}}
```

```
T13 = {T, 13};
Fs = {U, {.4, .9}};
Kp = {N, {.263, .018}};
cdfs = {Fmr, SABW, Fs, Kp, ETshower, TresAng, T13};
Clear[fxn];
fxn[f_, sabw_, fs_, kp_, etsh_, t_, t13_] :=
```

$$f * sabw * fs * kp * etsh * t \, \frac{EFcon}{ATcon} \, ConcAng \left(1 - \frac{.76 + .029 \, t13}{2}\right) 10^{-3}$$

```
fxn[FmrBar, SABWBar, .65, .263, .129, TresAngBar, 0]
```

6.92183×10^{-6}

Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, 2000, 10, Scale → 1000];]

{131.933 Second, Null}

TBL/@jen

{$\begin{matrix} \text{Mean}[\Delta r] & \text{Max}[|\Delta r|] & \text{JennrichChi2} & \text{DegFr} & \text{P-adj} \\ -0.000538521 & 0.0183593 & 2.62117 & 21 & 1. \end{matrix}$,

Fractile	Value	CVM(%)
0.01	0.0000165555	3.56479
0.05	0.0000796803	1.07369
0.5	0.00159787	0.732505
0.95	0.0331107	1.28518
0.99	0.0669827	1.42869
Mean	0.00684773	0.342052
Variance	0.000191493	1.74394

PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .03];
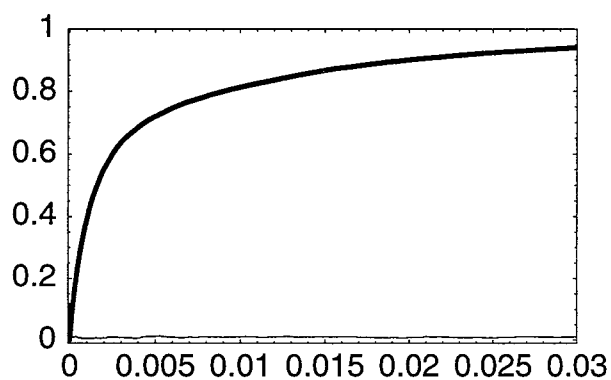
Put[cdf, PathName[EdermalAng]];

end

end

Appendix 3.E:
Effective Cytotoxic Dose

```
<< RiskQ`;

HardDrive = "Bogen";
PathName[filename_, hardDrive_String: HardDrive] := Module[{file = filename},
        If[Head[file] =!= String, file = ToString[file]];
        StringJoin[hardDrive, ":Ken:TCE Air Force:Data:", file]
    ];
```

Inputs

■ Log-Transform Utility Functions

```
GMGSDx::usage =
  "GMGSDx[Mx,SDx] returns the  geometric mean and geometric sdandard deviation of
    a lognormal variate X that also has the specified arithmetic mean Mx and
    arithmetic sdandard deviation SDx, based on the method of moments.";

MSDx::usage =
  "MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation
    of a lognormal variate X that also has the specified geometric mean GMx and
    geometric sdandard deviation GSDx, based on the method of moments.";

GMGSDx1::usage =
  "GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that
    the product X1*X2 has the desired coefficient of variation (CV) = cvWant,
    conditional on the lognormal variate X2 having an arithmetic mean and
    CV equal to 1 and cv2, respectively, based on the method of moments.";

MSDx[GMx_, GSDx_] := Module[{mux, sigy},
        sigy = Log[GSDx];
        mux = GMx E ^ ((sigy^2) / 2);
        mux {1, Sqrt[E ^ (sigy^2) - 1]}]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
        sigy = Sqrt[Log[1 + (SDx / Mx) ^ 2]];
        muy = Log[Mx] - (sigy^2) / 2;
        E ^ {muy, sigy}]

GMGSDx1[cvWant_, cv2_] := Module[{my1},
        muy1 = Log[Sqrt[(cv2^2 + 1) / (cvWant^2 + 1)]];
        E ^ {muy1, Sqrt[-2 muy1]}
    ] /; cvWant >= cv2
```

■ Input Empirical (Derived) Distributions

```
Conc = ToExpression[
   ReadList[PathName["concentration.txt", HardDrive], Word, RecordLists -> True]];
ConcAng =  0.0223;   (* mg/L    *)


Ingest = {LN, {Log[.0198], Log[1.88]}};
IngestBar =  0.0242; (* L/kg-d  *)


SABW = ToExpression[
   ReadList[PathName["sabwratioALL.txt", HardDrive], Word, RecordLists -> True]];
SABWBar =  325.881; (* cm2/kg *)
Fs = {U, {.4, .9}};
Kp = {N, {.263, .018}};


InhaleBar = 12.9 * (71^.74);   (* L/h *)
VQ = {LN, Log[{.959938, 1.33104}]};
T13 = {T, 13};
AEshower = {U, {4, 20}};
AEshHBar = 1 / EV[{1. / #[[1]], #[[2]]}& /@RQ[Cdf, U, {4, 20}, 2000]];
{gmgsd = GMGSDx@@#& /@N[{{480, 160}, {.129, .052}}],
{Wshower, ETshower} = ({LN, #}& /@Log[gmgsd])}

{{{455.368, 1.38347}, {0.119645, 1.47407}},
  {{LN, {6.12111, 0.324593}}, {LN, {-2.12323, 0.388026}}}}
```

end

■ Feq (function)

$$\text{lim} = \text{Limit}\left[\frac{1 - E^{-k\,t}}{t\,(1 - E^{-k\,24})},\, t \to 0\right]$$

$$\frac{k}{1 - E^{-24\,k}}$$

```
{limk = lim /. k -> .000000000001, 1 / limk}
```

```
{0.0416666, 24.}
```

$$fx[k\_,\, t\_] := \frac{1 - E^{-k\,t}}{t\,(1 - E^{-k\,24})}$$

```
Plot[{fx[k, .01], fx[k, .25], fx[k, .5], 24^-1 + .57 k}, {k, 0.0001, .1},
 AxesOrigin -> {0.0001, 0.04}, PlotRange -> {{0.0001, .1}, {0.04, .12}}];
```
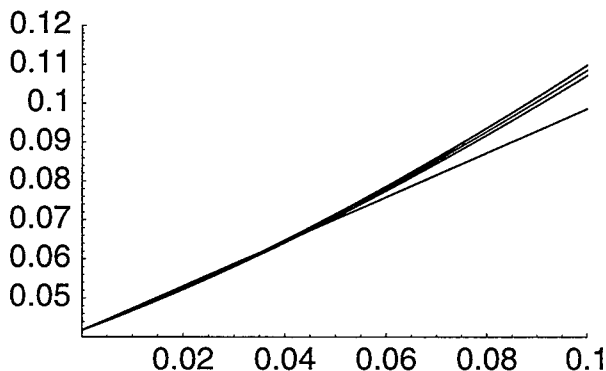


```
kval = Join[{.001}, Range[.005, .1, .0025]];
out = {fx[k, .01], fx[k, .25], fx[k, .5]} /. k -> kval;
xy = Flatten[Transpose[{kval, #}]& /@ (out - 24^-1), 1];
fit = Fit[xy, {x, x^2}, x] (* k = 0 - .1 *)
```

$$0.505316\,x + 1.66078\,x^2$$

```
fit[[2]] / fit[[1]]
```

$$3.28661\,x$$

```
zz = RQ[Q, N, {0, 1}, {.95, .99, .995}];
cv = 0.60; (* = assumed CV for Vke; see below *)
gsd = E^Sqrt[Log[1 + cv^2]]; (* by method of moments *)
{ke = 0.028 * 71^-.3, gsd, ke * gsd^zz}
```

$$\{0.00779436, 1.74109, \{0.0194043, 0.028315, 0.032516\}\}$$

```
data = ({0, 24^-1} + #)& /@ Prepend[xy, {0, 0}];
PlotData[data, FitTo -> {{24^-1 + fit, 24^-1 + 0.5053 x}, x},
Xmin -> -.0001, Xmax -> .1, Ymin -> .0399, Ymax -> .12, Style -> OO, Dashed -> {False, .025},
 DotSize -> .0125];
```
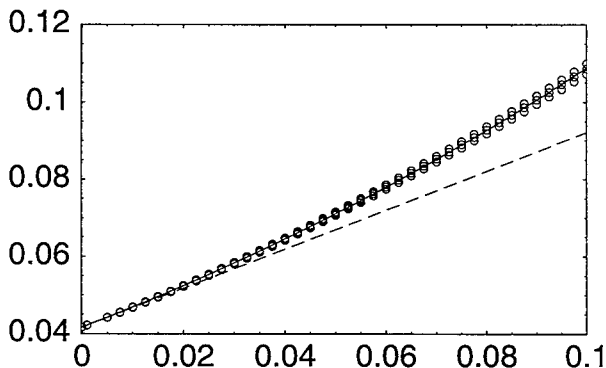
```
(* ~% relative error of linear approximation *)
Plot[100 ((24^-1+fit) / (24^-1+0.5053 x) -1), {x, 0, .06},
 AxesOrigin → {-.0001, -.0001}, PlotRange → {{-.0001, .06}, {-.0001, 10}},
 Frame → True, GridLines → {
    Range[.01, .05, .01], Range[2, 8, 2]}];
```
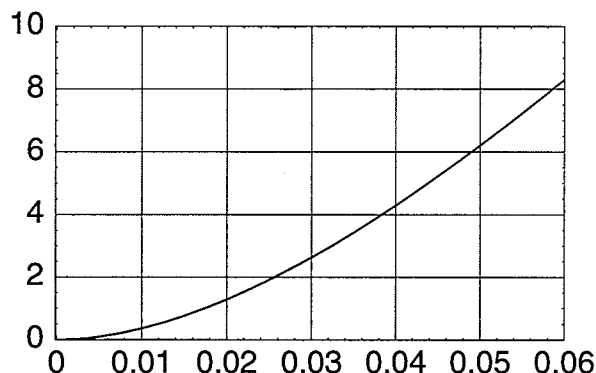

end

■ Body Weight (adult male + female), Vmax = Ve, and VQ

CV   = coefficient of variation

Qa   = alveolar ventilation rate = 12.9*BW^.74 (Allen & Fisher, 1993)

VVmax= Variability (unitless) in Vmax, where Vmax = 14.9*BW^.74 (in mg/h) (Allen & Fisher, 1993)

Vinhale= Variability (unitless) in Inhalation rate, where latter in L/h

VKe   = Variability (unitless) in Ke, where Ke = 0.028*BW^-.3 (in 1/h) (Allen & Fisher, 1993)

From Finley et al. 1994 (CalEPA 1996, p. 10-7), the BW distribution for adult males & females is ~LN and CV[BW] = ~0.22:

```
Clear[gsd];
{ev, sd} = {71., 15.9};
{cv = sd / ev, cvWant = 0.6, {BWgm, BWgsd} = GMGSDx[ev, sd],
Log[{BWgm, BWgsd}]}

{0.223944, 0.6, {69.2839, 1.24759}, {4.23821, 0.22121}}
```

CVwant = 0.60 for Vmax/BW assumed, based on Lipscombe et al. 1998 (Table 5). From Allen and Fisher (1993), the maximum metabolic rate in mg/h is modeled as Vmax ~ 14.9*BW^0.7. Now let VVmax be LN-distributed with an arithmetic mean of 1, where Vmet represents variation in Vmax not attributable to that in BW. Thus, Vmax ~ 14.9*Vmet*BW^.7, whence Vmax/BW ~ 14.9*Vmet*BW^-0.3. It follows from the method of moments that CV[BW^-.3] = CV[BW^.3] = 0.06644, whence CV[Vmet] = 0.5950, GM[Vmet] = 0.8594, GSD[Vmet] = 1.734, Log[{GM[Vmet], GSD[Vmet]}] = {-0.15154,0.550528}. CVwant = 0.60 for VKe is also assumed, based on Fisher et al. 1998 (Table 8); thus VKe = VVmax = Vmet.

```
{{o1, o2} = MSDx[a BWgm^-.3, BWgsd^-.3], cvCorr = o2 / o1}

{{0.281039 a, 0.0186711 a}, 0.066436}
```

```
{{o1, o2} = MSDx[a BWgm^-.3, BWgsd^.3], cvCorr = o2 / o1}
```

```
{{0.281039 a, 0.0186711 a}, 0.066436}
```

```
{o = GMGSDx1[cvWant, cvCorr], {muyVmet, sdyVmet} = Log[o]}
```

```
{{0.859383, 1.73417}, {-0.15154, 0.550528}}
```

```
MSDx@@o
```

```
{1., 0.594999}
```

```
(* By definition, the CV of (BW^.3 * Vmet) = *)
Sqrt[E^(sdyVmet^2 + Log[BWgsd^.3]^2) - 1]
```

```
0.6
```

```
{BW = {LN, Log[{BWgm, BWgsd}]}, Vmet = {LN, {muyVmet, sdyVmet}}}
```

```
{{LN, {4.23821, 0.22121}}, {LN, {-0.15154, 0.550528}}}
```

From CalEPA/OHEA (1996, Stochastic Analysis, p. 3-31 - 3-32; cit. above), cvA = CV[24*Q/BW] = CV[Qtot/BW] = ~ 0.3, where Q denotes total ventilation rate in L/h. From Allen and Fisher (1993), alveolar ventilation rate in L/h is modeled as Q ~ 12.9*BW^0.7, and Qtot ~ kQ for some constant k. Now let VQ be LN-distributed with an arithmetic mean of 1, where VQ represents variation in Q not attributable to that in BW. Thus, Q ~ 12.9*VQa*BW^.7, whence Q/BW ~ 12.9*VQ*BW^-0.3. It follows from the method of moments that CV[BW^-.3] = CV[BW^.3] = 0.06644, whence CV[VQ] = 0.2919, GM[VQ] = 0.9599, GSD[VQ] = 1.331, Log[{GM[VQ], GSD[VQ]}] = {-0.0408868, 0.285961}.

```
{{o1, o2} = MSDx[a BWgm^-.3, BWgsd^-.3], cvBW3 = o2 / o1}
```

```
{{0.281039 a, 0.0186711 a}, 0.066436}
```

```
{o = GMGSDx1[0.3, cvBW3], {muyX, sdyX} = Log[o]}
```

```
{{0.959938, 1.33104}, {-0.0408868, 0.285961}}
```

```
{{mX, sdX} = MSDx@@o, cvX = sdX / mX}
```

```
{{1., 0.291908}, 0.291908}
```

```
(* By definition, the CV of (BW^.3 * X) = *)
{Sqrt[E^(sdyX^2 + Log[BWgsd^.3]^2) - 1],
((1 + cvBW3^2) (1 + cvX^2) - 1)^.5}
```

```
{0.3, 0.3}
```

```
{BW,Vmet,VQ} =
   {{LN, {4.23821,0.22121}},
      {LN, {-0.15154,0.550528}},{LN, {-0.0408868,0.285961}}};
```

■ Fractions metabolized

■ Oral (fmo)

Heterogeneous variates ={Pb,Vmet,VQ}

```
{nsam, nsim} = {2000, 10};
Pb = NormalCdf[10.2, 1.6, 405];
cdfs = {Pb, VQ, Vmet};
Clear[fxn];
fxn[pb_, vq_, vmet_] := (1 + vq (vmet (.77 pb + 2.547))^-1)^-1

fxn[10.2, 1, 1]

0.912288

sim = Table[SimulateCdf[cdfs, 500, TestCdf → False, Report → Append], {10}];
{sims, rval, jens} = Transpose[sim];
jen = Last[Sort[Last /@ jens]]   (* Max[chi2],df,pval *)

{0.398663, 3, 0.940519}

corr = First[Correlation[{#[[3]], fxn@@#& /@Transpose[#]},
   Type → Spearman, Report → False]]& /@sims;
Stats[corr, Report]
```

Mean	SD	CVM%	95%LCL	95%UCL	Min	Max	n
0.859935	0.00411002	0.1511	0.856994	0.862875	0.854597	0.868392	10

Correlation between Vmet and Fmo = 0.86

```
corr = First[Correlation[{#[[2]], fxn@@#& /@Transpose[#]},
   Type → Spearman, Report → False]]& /@sims;
Stats[corr, Report]
```

Mean	SD	CVM%	95%LCL	95%UCL	Min	Max	n
−0.43378	0.0133479	0.9731	−0.443329	−0.424232	−0.45298	−0.417722	10

Correlation between VQ and Fmo = -0.43

$$corr = First\left[Correlation\left[\left\{\left(1 + \frac{\#[[2]]}{\#[[1]]} \left(\frac{1.299}{\#[[3]]} + 3.307\right)\right)^{-1}, fxn@@#\& /@Transpose[#]\right\},\right.\right.$$
Type → Spearman, Report → False]]& /@sims;
Stats[corr, Report]

Mean	SD	CVM%	95%LCL	95%UCL	Min	Max	n
0.825436	0.00851143	0.3261	0.819347	0.831524	0.81263	0.837117	10

Correlation between Fmr and Fmo = 0.83

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1];]

{1094.55 Second, Null}
```

TBL/@jen

$\left\{ \begin{array}{llll} \text{Mean}[\Delta r] & \text{Max}[|\Delta r|] & \text{JennrichChi2} & \text{DegFr} \quad \text{Pval} \\ 0.00256474 & 0.0100917 & 0.390451 & 3 \qquad 0.942208 \end{array} \right\}$

```
Fractile  Value       CVM(%)
0.01      0.680326    0.395848
0.05      0.765428    0.103074
0.5       0.90237     0.0203722  }
0.95      0.963376    0.0237576  }
0.99      0.975746    0.0236983
Mean      0.888543    0.00239931
Variance  0.00395071  0.863083
```

{First[cdf], Last[cdf]}

{{0.476647, 0}, {0.99139, 1}}

PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → .495, Xmax → 1];
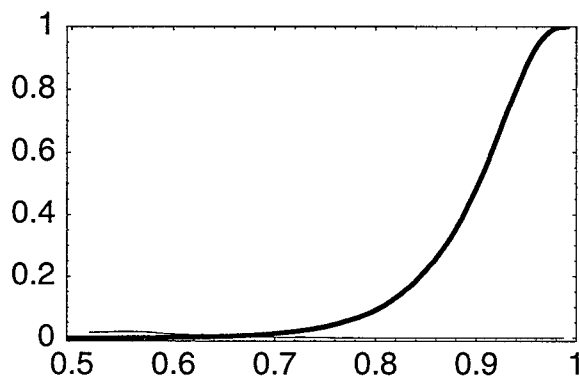
Fmo = StandardizeCdf[cdf, 405];

end

■ Inhalation and dermal (fmr)

Heterogeneous variates ={Pb,Vmet,VQ}

{nsam, nsim} = {2000, 10};
cdfs = {Pb, VQ, Vmet};
Clear[fxn];

$$\text{fxn}[pb\_, vq\_, vmet\_] := \left(1 + \frac{vq}{pb}\left(\frac{1.299}{vmet} + 3.307\right)\right)^{-1}$$

fxn[10.2, 1, 1]

0.68891

corr = First[Correlation[{#[[2]], fxn@@#&/@Transpose[#]},
  Type → Spearman, Report → False]]&/@sims;
Stats[corr, Report]

Mean	SD	CVM%	95%LCL	95%UCL	Min	Max	n
-0.753049	0.00720099	0.3024	-0.758201	-0.747898	-0.764949	-0.741702	10

Correlation between VQ and Fmr = -0.75

```
corr = First[Correlation[{#[[3]], fxn@@#&/@Transpose[#]},
   Type → Spearman, Report → False]]&/@sims;
Stats[corr, Report]
```

```
Mean      SD          CVM%     95%LCL    95%UCL    Min       Max        n
0.453276  0.00968226  0.6755   0.44635   0.460202  0.436792  0.466869   10
```

Correlation between Vmet and Fmr = 0.45

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1];]
```

```
{1096.72 Second, Null}
```

TBL/@jen

					Fractile	Value	CVM(%)
					0.01	0.454963	0.559908
					0.05	0.527838	0.151762
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	Pval	0.5	0.680315	0.0680913
0.001593	0.0140407	0.499243	3	0.919058	0.95	0.794697	0.0632056
					0.99	0.830531	0.090292
					Mean	0.673284	0.0011681
					Variance	0.00656651	0.240044

{First[cdf], Last[cdf]}

```
{{0.264383, 0}, {0.890069, 1}}
```

PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → 1];
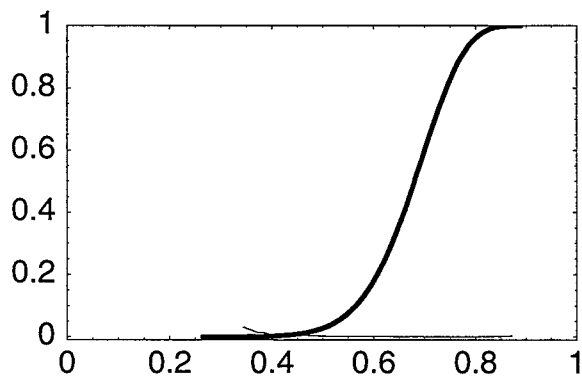
Fmr = StandardizeCdf[cdf, 405];

end

Put[Fmo, PathName["Fmo"]];

Put[Fmr, PathName["Fmr"]];

Correlation between Vmet and Fmo = 0.86
Correlation between VQ and Fmr = -0.75
Correlation between Vmet and Fmr = 0.45

```
Fmo = Get[PathName[Fmo]]; Fmr = Get[PathName[Fmr]];
CdfQ/@{Fmo, Fmr}
```

{True, True}

{FmoBar, FmrBar} = {0.888543, 0.6732836};

```
PlotCdf[{Fmr, Fmo}, Ymin → -.001, Xmin → .398, Xmax → 1];
```
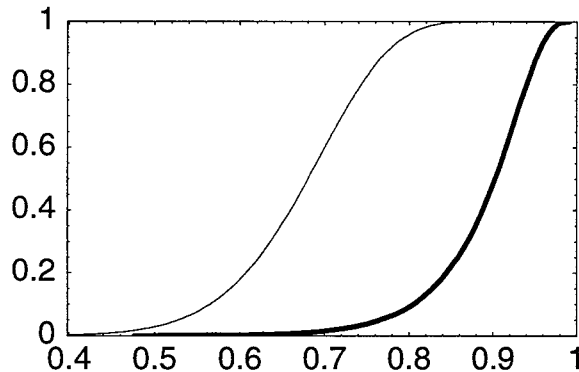

end

end

Note: All distributions below are *un*scaled

■ Ingestion Effective Dose (mg TCA/L plasma)

■ <ECingest> = Variability in Expected Level

{VolDist,BW,VKe,Fmo,Ingest} = heterogeneous variates
Correlation between VolDist and BW is assumed to be -.5
Correlation between Vmet =Ve and Fmo is assumed to be 0

```
{nsam, nsim} = {2000, 10};
Ingest = {LN, {Log[.0198], Log[1.88]}};
VolDist = {U, {.052, .152}}; (* L/kg *)
cdfs = {VolDist, BW, Vmet, Fmo, Ingest};
corr = {{-.5, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
Clear[fxn];
```

$$fxn[u\_, bw\_, vmet\_, fmo\_, ing\_] := ing * ConcAng * fmo * \left(\frac{.4104}{u} \left(\frac{1.488\ bw^{-3}}{vmet} + .5053 \right) \right)$$

```
fxn[.1, 70, 1, .7, IngestBar]
```

0.00903536

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1, Correlate → corr];]
```

{85.7333 Second, Null}

```
TBL/@jen
```

					Fractile	Value	CVM(%)
					0.01	0.00149497	1.41686
					0.05	0.00265397	0.846307
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj	0.5	0.0110646	0.571085
−0.000233837	0.0171728	1.5172	10	1.	0.95	0.0476501	0.587508
					0.99	0.0879985	2.44676
					Mean	0.0163479	0.171948
					Variance	0.000306492	3.01902

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .07];
```
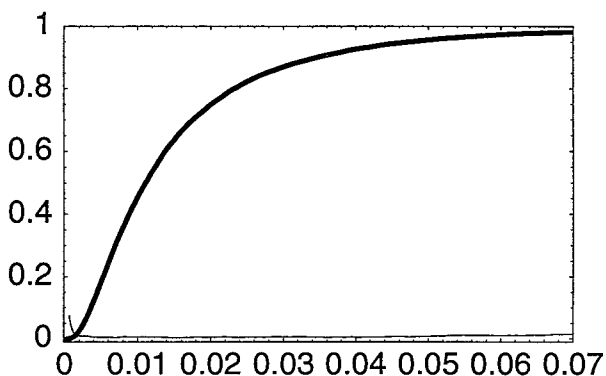


```
sdf = StandardizeCdf[cdf, 404]; EV[sdf, Empirical -> True]
```

0.0161848

```
Put[sdf, PathName[ECingestAng]];
```

```
ECingestAng = Get[PathName[ECingestAng]];
```

```
ECingestAngBar=0.0161848;
```

end

▪ ECingestBar = Uncertainty in Population-Average Level

{Conc} = uncertain variate

$$cdf = \left\{ \frac{ECingestAngBar}{ConcAng} \, \#[[1]], \, \#[[2]] \right\} \& /@Conc;$$

```
Put[cdf, PathName[ECingestBar]];
```

```
ECingestBar = Get[PathName[ECingestBar]];
```

end

end

■ Inhalation Effective Dose (mg TCA/L plasma)

■ <ECinhale> = Variability in Expected Level

{VolDist,BW,VKe,Fmr,Inhale,Wshower,ETshower,AEshower,T13} = heterogeneous variates
Correlation between Vmet = Ve and Fmr is assumed to be 0
Correlation between VolDist and BW is assumed to be -.5
Correlation between VQ and Fmr = -0.75

```
{InhaleBar, ConcAng}

{302.358, 0.0223}

{nsam, nsim} = {2000, 10};
Clear[fxn];
cdfs = {VolDist, BW, Vmet, VQ, Fmr, Wshower, ETshower, AEshower, T13};
corr = {{-.5, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0},
   {0, 0, 0, 0, 0, 0}, {-.75, 0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
```

$$fxn[u\_, bw\_, vmet\_, vq\_, fmr\_, wsh\_, etsh\_, aesh\_, t13\_] := \left(\frac{.4104}{u}\left(\frac{19.195}{vmet} + \frac{6.5145}{bw^{.3}}\right)\right.$$

$$\left.\right)vq * ConcAng\left(\frac{wsh\,(.76 + .029\,t13)}{1000\,aesh}\right)fmr * etsh$$

```
fxn[.1, 70, 1, 1, .7, 480, .129, AEshHBar, 0]

0.00637328

Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1, Correlate → corr];]

{179.167 Second, Null}

TBL/@jen
```

Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj
-0.000248172	0.0207595	4.33342	36	1.

Fractile	Value	CVM(%)
0.01	0.000715826	1.62597
0.05	0.00126666	0.60722
0.5	0.00538319	0.46861
0.95	0.0259794	1.09888
0.99	0.0489231	1.70975
Mean	0.00844442	0.179915
Variance	0.0000944227	3.53902

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → .05];
```
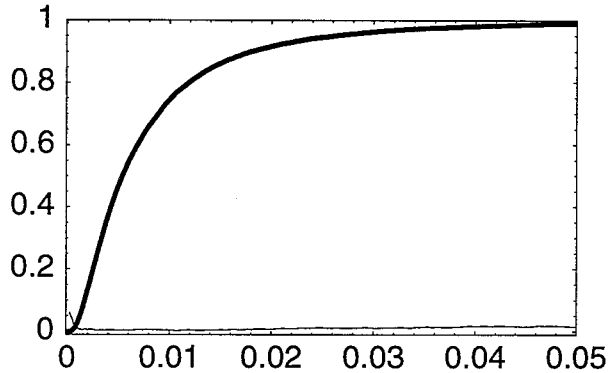


```
sdf = StandardizeCdf[cdf, 404]; EV[sdf, Empirical → True]

0.00835208

Put[sdf, PathName[ECinhaleAng]];

ECinhaleAng = Get[PathName[ECinhaleAng]];

ECinhaleAngBar = 0.0083520792794854345`;
```

end

■ ECinhaleBar = Uncertainty in Population-Average Level

{Conc} = uncertain variate

$$cdf = \left\{ \frac{ECinhaleAngBar}{ConcAng} \#[[1]], \#[[2]] \right\} \text{\&/@Conc;}$$

```
Put[cdf, PathName[ECinhaleBar]];

ECinhaleBar = Get[PathName[ECingestBar]];
```

end

end

■ Dermal Effective Dose (mg TCA/L plasma)

■ <Edermal> = Variability in Expected Level

{VolDist,BW,Vmet,Fmr,SABW,Fs,Kp,ETshower,T13} = heterogeneous variates
Correlation between Vmet = Ve and Fmr is assumed to be 0
Correlation between VolDist and BW is assumed to be -.5

```
{nsam, nsim} = {2000, 10};
corr = {{-.5, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0},
   {0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
cdfs = {VolDist, BW, Vmet, Fmr, SABW, Fs, Kp, ETshower, T13};
Clear[fxn];
fxn[u_, bw_, vmet_, fmr_, sabw_, fs_, kp_, etsh_, t13_] :=
```

$$\text{sabw} * \text{fs} * \text{kp} * \text{etsh} * \text{ConcAng}\left(1 - \frac{.76 + .029\, t13}{2}\right) 10^{-3} * \text{fmr} * \left(\right.$$

$$\left.\frac{.4104}{u}\left(\frac{1.488\, bw^{.3}}{vmet} + .5053\right)\right)$$

```
fxn[.1, 70, 1, .7, SABWBar, .65, .263, .129, 0]
```

```
0.00166356
```

```
(CdfQ[#] || RQ[Test, #[[1]], #[[2]]])&/@cdfs
```

```
{True, True, True, True, True, True, True, True, True}
```

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Scale → 1, Correlate → corr];]
```

```
{173.433 Second, Null}
```

```
TBL/@jen
```

						Fractile	Value	CVM(%)
						0.01	0.000268902	0.97206
						0.05	0.000445612	0.780994
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj		0.5	0.00163022	0.369624
0.000459795	0.0212962	4.44316	36	1.		0.95	0.00625184	1.02696
						0.99	0.0111429	0.914124
						Mean	0.00228462	0.1228
						Variance	5.0004×10^{-6}	2.93471

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.00001, Xmax → .01];
```
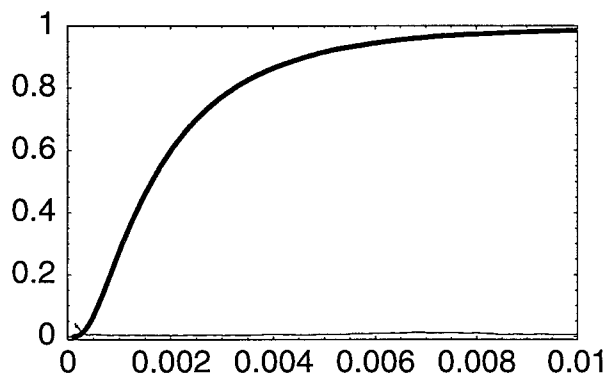


```
sdf = StandardizeCdf[cdf, 404]; EV[sdf, Empirical → True]
```

```
0.00226194
```

```
Put[sdf, PathName[ECdermalAng]];
```

```
ECdermalAng = Get[PathName[ECdermalAng]];
```

```
ECdermalAngBar = 0.00226194038376105455`;
```

end

■ EdermalBar = Uncertainty in Population-Average Level

{Conc} = uncertain variates

$$cdf = \left\{ \frac{\texttt{ECdermalAngBar}}{\texttt{ConcAng}} \, \#[[1]], \, \#[[2]] \right\} \& \, /@Conc;$$

```
Put[cdf, PathName[ECdermalBar]];
```

```
ECdermalBar = Get[PathName[ECdermalBar]];
```

end

end

Appendix 3.F

Effective Dose Correlations

```
<< RiskQ`;
HardDrive = "Bogen";
PathName[filename_, hardDrive_String: HardDrive] := Module[{file = filename},
        If[Head[file] =!= String, file = ToString[file]];
        StringJoin[hardDrive, ":Ken:TCE Air Force:Data:", file]
    ];
```

Inputs

■ Log-Transform Utility Functions

```
GMGSDx::usage =
  "GMGSDx[Mx,SDx] returns the  geometric mean and geometric sdandard deviation of
    a lognormal variate X that also has the specified arithmetic mean Mx and
    arithmetic sdandard deviation SDx, based on the method of moments.";

MSDx::usage =
  "MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation
    of a lognormal variate X that also has the specified geometric mean GMx and
    geometric sdandard deviation GSDx, based on the method of moments.";

GMGSDx1::usage =
  "GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that
    the product X1*X2 has the desired coefficient of variation (CV) = cvWant,
    conditional on the lognormal variate X2 having an arithmetic mean and
    CV equal to 1 and cv2, respectively, based on the method of moments.";

MSDx[GMx_, GSDx_] := Module[{mux, sigy},
        sigy = Log[GSDx];
        mux = GMx E ^ ((sigy^2) / 2);
        mux {1, Sqrt[E ^ (sigy^2) - 1]}]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
        sigy = Sqrt[Log[1 + (SDx / Mx) ^ 2]];
        muy = Log[Mx] - (sigy^2) / 2;
        E ^ {muy, sigy}]

GMGSDx1[cvWant_, cv2_] := Module[{my1},
        muy1 = Log[Sqrt[ (cv2^2 + 1) / (cvWant^2 + 1) ]];
        E ^ {muy1, Sqrt[-2 muy1]}
    ] /; cvWant >= cv2
```

■ Input Empirical (Derived) Distributions

```
Clear[Tbarang, TresBar, TresAng];
Tbarang = Rest[Get[PathName[Tbarang]]];
TresBar = #[[{1, 2}]]&/@Tbarang;
TresAng = #[[{1, 3}]]&/@Tbarang;
TresBarAng =  7.59358; (* y *)
TresAngBar =  7.55321; (* y *)
EFcon  =   350;  (* d/y  *)
ATcon  = 25550;  (* d    *)
Conc = ToExpression[
   ReadList[PathName["concentration.txt", HardDrive], Word, RecordLists -> True]];
ConcAng =  0.0223;  (* mg/L   *)


Ingest = {LN, {Log[.0198], Log[1.88]}};
IngestBar =  0.0242; (* L/kg-d  *)
ECingestAngBar = 0.0161848;


SABW = ToExpression[
   ReadList[PathName["sabwratioALL.txt", HardDrive], Word, RecordLists -> True]];
SABWBar  =   325.881; (* cm2/kg *)
Fs = {U, {.4, .9}};
Kp = {N, {.263, .018}};
ECdermalAngBar = 0.002261940;


Inhale = ToExpression[ReadList[PathName["inhaleALL.txt"], Word, RecordLists -> True]];
     (*Note: Inhale in L/kg-d *)
InhaleBarC = 12.9*(71^.74);   (* L/h *)
InhaleBarG = 0.102;   (* m3/kg-d *)
ECinhaleAngBar = 0.0083520793;
{BW, Vmet, VQ} =
 {{LN, {4.23821, 0.22121}},
  {LN, {-0.15154, 0.550528}}, {LN, {-0.0408868, 0.285961}}};
VolDist = {U, {.052, .152}}; (* L/kg *)
T13 = {T, 13};


{AEshHBar, AEbaHBar, AEhHBar} = 1 / (EV/@{
{1./#[[1]], #[[2]]}&/@RQ[Cdf, U, {4, 20}, 2000],
{1./#[[1]], #[[2]]}&/@RQ[Cdf, U, {10, 100}, 2000],
{1./#[[1]], #[[2]]}&/@RQ[Cdf, U, {300, 1200}, 2000]});
gmgsd = GMGSDx@@#&/@N[{{480, 160}, {42, 15}, {.129, .052}, {.33, .22}}];
{Wshower, Whouse, ETshower, ETbath} = ({LN, #}&/@Log[gmgsd]);
{AEshower, AEbath, AEhouse, EThouse} =
 ({U, #}&/@{{4, 20}, {10, 100}, {300, 1200}, {8, 20}});
```

end

■ Fractions metabolized (summary–see Effective Cytotoxic Dose.nb)

Correlation between VolDist and BW = -0.50 (assumed approximation)
Correlation between Ve and Fmo =0, and between Ve and Fmr = 0
Correlation between VQ and Fmo = -0.43
Correlation between VQ and Fmr = -0.75
Correlation between Fmr and Fmo = 0.83

```
{FmoBar, FmrBar} = {0.888543, 0.6732836};
Clear[Fmo, Fmr];
Fmo = Get[PathName[Fmo]]; Fmr = Get[PathName[Fmr]];
CdfQ/@{Fmo, Fmr}

{True, True}
```

end

end

■ Effective Dose Uncertainty

```
cdfGingBar = {TresBar, Conc};
```
$$\text{fxnGingBar}[t\_, c\_] := \text{IngestBar} * t \; \frac{\text{EFcon}}{\text{ATcon}} \; c * \text{FmoBar};$$
```
cdfCingBar = {Conc};
```
$$\text{fxnCingBar}[c\_] := c \; \frac{\text{ECingestAngBar}}{\text{ConcAng}};$$
```
cdfGinhBar = {TresBar, Conc};
```
$$\text{fxnGinhBar}[t\_, c\_] := \text{FmrBar} \left(\frac{12.9}{1000} * 71^{.74-1} * 1\right) * t$$
$$\frac{\text{EFcon}}{\text{ATcon}} \; c \; \frac{1}{24} \left(480 \,(.76) \left(\frac{.129}{\text{AEshHBar}} + \frac{.33}{\text{AEbaHBar}}\right) + 42 \left(.76 \; \frac{.54}{.7}\right) \frac{14}{\text{AEhHBar}}\right);$$
```
cdfCinhBar = {Conc};
```
$$\text{fxnCinhBar}[c\_] := c \; \frac{\text{ECinhaleAngBar}}{\text{ConcAng}};$$
```
cdfGderBar = {TresBar, Conc};
```
$$\text{fxnGderBar}[t\_, c\_] := \text{FmrBar} * \text{SABWBar} * .65 * .263 * .129 * t \; \frac{\text{EFcon}}{\text{ATcon}} \; c \left(1 - \frac{.76}{2}\right) 10^{-3};$$
```
cdfCderBar = {Conc};
```
$$\text{fxnCderBar}[c\_] := c \; \frac{\text{ECdermalAngBar}}{\text{ConcAng}};$$

All 6 functions above are linear functions of either c or c*t; thus: all those involving just c are 100% correlated, all those involving just c*t are 100% correlated, and correlations between those involving c vs. c*t are given by:

```
cdfs = {TresBar, Conc};
sim = Table[SimulateCdf[cdfs, 500, TestCdf → False, Report → Append], {10}];
{sims, rval, jens} = Transpose[sim];
jen = Last[Sort[Last/@jens]]   (* Max[chi2],df,pval *)

{0.339124, 1, 0.560336}


corr = First[Correlation[{#[[2]], #[[1]] * #[[2]]},
    Type → Spearman, Report → False]]&/@sims;
Stats[corr, Report]

Mean       SD         CVM%     95%LCL    95%UCL   Min       Max       n
0.487765   0.0103941  0.6739   0.480329  0.4952   0.470525  0.506052  10
```

end

■ Effective Dose Variability

```
cdfGingAng = {Ingest, TresAng, Fmo};
```

$$\text{fxnGingAng}[ing\_, t\_, f\_] := ing * t \, \frac{\text{EFcon}}{\text{ATcon}} \, \text{ConcAng} * f;$$

```
cdfCingAng = {VolDist, BW, Vmet, Fmo, Ingest};
corCingAng = {{-.5, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
```

$$\text{fxnCingAng}[u\_, bw\_, vmet\_, fmo\_, ing\_] := ing * \text{ConcAng} * fmo * \left(\right.$$

$$\frac{.4104}{u} \left(\frac{1.488\, bw^{.3}}{vmet} + .5053 \right) \Bigg) \Bigg);$$

```
cdfGinhAng = {Fmr, VQ, BW, TresAng, Wshower,
    Whouse, ETshower, ETbath, EThouse, AEshower, AEbath, AEhouse, T13, T13};
corr = Table[Table[0, {j}], {j, 13}];
corGinhAng = ReplacePart[Reverse[corr], -.75, {1, 1}]; fxnGinhAng[f_, vq_,
    bw_, t_, wsh_, wh_, etsh_, etba_, eth_, aesh_, aeba_, aeh_, t13sh_, t13h_] :=
```

$$f * \left(\frac{12.9}{1000} * bw^{.74-1} * vq \right) t \, \frac{\text{EFcon}}{\text{ATcon}} \, \text{ConcAng} \, \frac{1}{24} \left(\right.$$

$$wsh\,(.76 + .029\, t13sh) \left(\frac{etsh}{aesh} + \frac{etba}{aeba} \right) + wh \left((.76 + .029\, t13h) \, \frac{.54}{.7} \right) \frac{eth}{aeh} \right);$$

```
cdfCinhAng = {VolDist, BW, Vmet, VQ, Fmr, Wshower, ETshower, AEshower, T13};
corCinhAng = {{-.5, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0},
    {0, 0, 0, 0, 0, 0}, {-.75, 0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
fxnCinhAng[
```

$$u\_, bw\_, vmet\_, vq\_, fmr\_, wsh\_, etsh\_, aesh\_, t13\_] := \left(\frac{.4104}{u} \left(\frac{19.195}{vmet} + \frac{6.5145}{bw^{.3}} \right) \right)$$

$$\Bigg) vq * \text{ConcAng} \left(\frac{wsh\,(.76 + .029\, t13)}{1000\, aesh} \right) fmr * etsh;$$

```
cdfGderAng = {Fmr, SABW, Fs, Kp, ETshower, TresAng, T13};
fxnGderAng[f_, sabw_, fs_, kp_, etsh_, t_, t13_] :=
```

$$f * sabw * fs * kp * etsh * t \, \frac{\text{EFcon}}{\text{ATcon}} \, \text{ConcAng} \left(1 - \frac{.76 + .029\, t13}{2} \right) 10^{-3};$$

```
cdfCderAng = {VolDist, BW, Vmet, Fmr, SABW, Fs, Kp, ETshower, T13};
corCderAng = {{-.5, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0},
    {0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
fxnCderAng[u_, bw_, vmet_, fmr_, sabw_, fs_, kp_, etsh_, t13_] :=
```

$$sabw * fs * kp * etsh * \text{ConcAng} \left(1 - \frac{.76 + .029\, t13}{2} \right) 10^{-3} * fmr * \left(\right.$$

$$\frac{.4104}{u} \left(\frac{1.488\, bw^{.3}}{vmet} + .5053 \right) \Bigg) \Bigg);$$

Function {argument positions}:
fxnGingAng[ing_,t_,f_]
     {7,8,5}
fxnCingAng[u_,bw_,vmet_,fmo_,ing_]
     {1,2,3,5,7}

fxnGinhAng[f_,vq_,bw_,t_,wsh_,wh_,etsh_,etba_,eth_,aesh_,aeba_,aeh_,t13sh_,t13h_]
      {6,4,2,8,9,10,11,12,13,14,15,16,17,18}
fxnCinhAng[u_,bw_,vmet_,vq_,fmr_,wsh_,etsh_,aesh_,t13_]
      {1,2,3,4,6,9,11,14,17}
fxnGderAng[f_,sabw_,fs_,kp_,etsh_,t_,t13_]
      {6,19,20,21,11,8,17}
fxnCderAng[u_,bw_,vmet_,fmr_,sabw_,fs_,kp_,etsh_,t13_]
      {1,2,3,6,19,20,21,11,17}

```
cdfs = {VolDist, BW, Vmet, VQ, Fmo, Fmr, Ingest, TresAng,
   Wshower, Whouse, ETshower, ETbath, EThouse, AEshower,
   AEbath, AEhouse, T13, T13, SABW, Fs, Kp}; (* n=21 *)

(CdfQ[#] || RQ[Test, #[[1]], #[[2]]])&/@cdfs

{True, True, True, True, True, True, True, True, True, True, True, True, True,
   True, True, True, True, True, True, True}
```

Correlation between VolDist and BW = -0.50 (assumed approximation)
Correlation between Ve and Fmo =0, and between Ve and Fmr = 0
Correlation between VQ and Fmo = -0.43
Correlation between VQ and Fmr = -0.75
Correlation between Fmr and Fmo = 0.83

```
corr = Reverse[Table[Table[0, {j}], {j, 20}]];
(corr = ReplacePart[corr, #[[1]], #[[2]]])&/@{
  {-.5, {1, 1}}, {-.43, {4, 1}}, {-.75, {4, 2}}, {.83, {5, 1}}};

Clear[o];
xx = {a, b, c, d, e, f, g,
   h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z}; xx

{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z}
```

Define fxns here each as a function of elements of the convenient dummy variate xx:

```
fxns = {
{fxnGingAng, {7, 8, 5}},
{fxnCingAng, {1, 2, 3, 5, 7}},
{fxnGinhAng, {6, 4, 2, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18}},
{fxnCinhAng, {1, 2, 3, 4, 6, 9, 11, 14, 17}},
{fxnGderAng, {6, 19, 20, 21, 11, 8, 17}},
{fxnCderAng, {1, 2, 3, 6, 19, 20, 21, 11, 17}}};
MapThread[#1@@xx[[#2]]&, Transpose[fxns]]
```

$$\left\{ 0.000305479\, e\, g\, h,\quad \frac{0.00915192\left(0.5053 + \frac{1.488\, b^{0.3}}{c}\right) e\, g}{a}, \right.$$

$$\frac{1.64195 \times 10^{-7}\, d\, f\, h\, (i\, (\frac{k}{n} + \frac{1}{o})\, (0.76 + 0.029\, q) + \frac{0.771429\, j\, m\, (0.76 + 0.029\, r)}{p})}{b^{0.26}},$$

$$\frac{9.15192 \times 10^{-6}\, (\frac{6.5145}{b^{0.3}} + \frac{19.195}{c})\, d\, f\, i\, k\, (0.76 + 0.029\, q)}{a\, n},$$

$$3.05479 \times 10^{-7}\, f\, h\, k\, \left(1 + \frac{1}{2}\, (-0.76 - 0.029\, q)\right) s\, t\, u,$$

$$\left. \frac{9.15192 \times 10^{-6}\, \left(0.5053 + \frac{1.488\, b^{0.3}}{c}\right) f\, k\, (1 + \frac{1}{2}\, (-0.76 - 0.029\, q))\, s\, t\, u}{a} \right\}$$

Select a series of 50 simulated-variate sets that each have a 21-variate correlation matrix (=corr, defined above) that is not significantly different than corr. Below, 20 of a total of 70 sets tried qualify for use using a p-value of <0.01 to reject:

```
Timing[
  sim = Table[(Prn[i]; SimulateCdf[cdfs, 500, Correlate → corr, Report → Append]), {i, 70}];
  {sims, rval, jens} = Transpose[sim];
  jen = Last[Sort[Last /@ jens]]   (* Max[chi2],df,pval *)]
```

{3335. Second, {35.2754, 210, 1.}}

```
Dimensions[sims]
```

{70, 21, 500}

```
sel = Select[o = Last /@ jens, Last[#] > 0.01&];
{pos = Position[o, #][[1, 1]]& /@ sel, Length[pos]}
```

{{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
  27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49,
  50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70},
 70}

```
Union[Last /@ Last /@ jens]
```

{1.}

```
OKsims = sims[[Take[pos, 50]]]; nn = Length[OKsims]
```

50

```
corm = Table[First[Correlation[
    MapThread[#1 @@ OKsims[[i]][[#2]]&, Transpose[fxns]],
  Type → Spearman, Report → False]], {i, nn}];
```

Mean corr, SDM[corr], and CVM[corr] values for:

      GingAng   **CingAng**   **GinhAng**   **CinhAng**   **GderAng**   **CderAng**

```
ev = Plus @@ corm / nn; TBL[N[ev, 3]]
```

1.	0.233	0.878	-0.00269	0.894	0.000496
0.233	1.	-0.0088	0.421	0.00439	0.514
0.878	-0.0088	1.	0.187	0.918	0.035
-0.00269	0.421	0.187	1.	0.0766	0.649
0.894	0.00439	0.918	0.0766	1.	0.177
0.000496	0.514	0.035	0.649	0.177	1.

```
sdm = √((Plus @@ (((# - ev)^2)& /@ corm)) / (nn (nn - 1))) ; TBL[N[sdm, 3]]
```

0	0.0022	0.000722	0.00207	0.000508	0.0018
0.0022	0	0.00207	0.00183	0.00223	0.00171
0.000722	0.00207	0	0.0022	0.000564	0.00192
0.00207	0.00183	0.0022	0	0.00205	0.00159
0.000508	0.00223	0.000564	0.00205	0	0.00214
0.0018	0.00171	0.00192	0.00159	0.00214	0

```
cvm = Abs[100 sdm / ev];  TBL[N[cvm, 3]]

0        0.946  0.0822  76.7    0.0568  364.
0.946    0      23.5    0.435   50.8    0.333
0.0822   23.5   0       1.17    0.0614  5.49
76.7     0.435  1.17    0       2.68    0.245
0.0568   50.8   0.0614  2.68    0       1.21
364.     0.333  5.49    0.245   1.21    0
```

end

Appendix 3.G

Potency

```
<<RiskQ`;

<<Minimize`;
```

Multistage (Genotoxicity) Model

■ Multistage Potencies for TCE Cancer Bioassays

The *Mathematica* program "QFit" (by K.T. Bogen, LLNL—see "RiskQ Functions Used" section below) was used to obtain for each bioassay data set a distribution reflecting parameter-estimation uncertainty pertaining to the value of multistage-model "potency" (denoted q_1), that is, the value of the linear coefficient of dose D in the multistage model of cancer risk, which posits that cancer risk is essentially an exponentiated-polynomial function dose. Conditional on any sufficiently "upper-bound" (i.e., conservative) estimate (denoted $q_1^* > 0$) of the linear "potency" term (q_1), the multistage model guarantees that any small increase in cancer risk will be very nearly equal to the product: $q_1^* \times D$. Uncertainty distributions are derived corresponding to each of seven bioassay data sets considered below; one data set (data set #7 below concerning the study by Henchler et al., 1980, showing malignant lymphoma in female HAN:NMRI mice) is excluded for reasons noted below.

■ 1. NCI 1976 Mouse B6C3F1: M 34 g   HCC

```
xhi=800;
     doses = { 0, 370, 739}*1. (* mg/kg-d LTWAM *);
   ndosed = {20,  48,  40};
 nrespond = { 1,  26,  31};
```

```
qTCE1=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->2, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

The Optimized Function F of Dose d is:

$F(d) = 1 - \exp[-P(d)]$, where:

$P(d) = 0.0524116 + 0.00199086\,d + 2.72013 \times 10^{-8}\,d^2$

ChiSquare(1)= 0.032011 1-tailed p= 0.858003

$R^2 = 0.999488$

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)

■ 2. NCI 1976 Mouse B6C3F1: F 29 g   HCC

```
xhi=600;
    doses = { 0, 275, 550}*1. (* mg/kg-d LTWAM *);
  ndosed = {18, 42, 37};
nrespond = { 0,  4, 11};
```

```
qTCE2=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->2, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

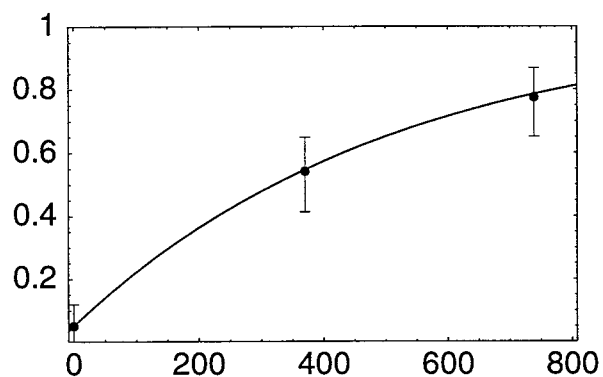The Optimized Function F of Dose d is:

$F(d) = 1 - \exp[-P(d)]$, where:

$P(d) = 0.0000863863\,d + 1.00929 \times 10^{-6}\,d^2$

ChiSquare(1) = 0. Perfect fit.

$R^2 = 1$.

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)

■ 3. NTP 1983 Mouse B6C3F1: M 37 g   HCC or HCA

```
xhi=600;
    doses = { 0, 563}*1. (* mg/kg-d LTWAM *);
  ndosed = {48, 50};
nrespond = {11, 38};
```

```
qTCE3=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->1, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

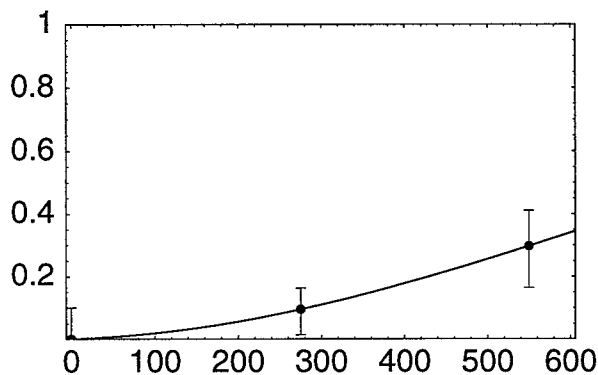The Optimized Function F of Dose d is:

F(d) = 1 - exp[-P(d)], where:

P(d) = 0.260283 + 0.00207253 d

ChiSquare(0) = 0. Perfect fit.

$R^2 = 1.$

```
            F(d) vs Data

    (& Bootstrap 90% Conf. Limits on Data)
```


■ 4. NTP 1983 Mouse B6C3F1: F 33 g   HCC or HCA:

```
xhi=600;
    doses = { 0, 563}*1. (* mg/kg-d LTWAM *);
  ndosed = {41, 41};
nrespond = { 4, 19};
```

```
qTCE4=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->1, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

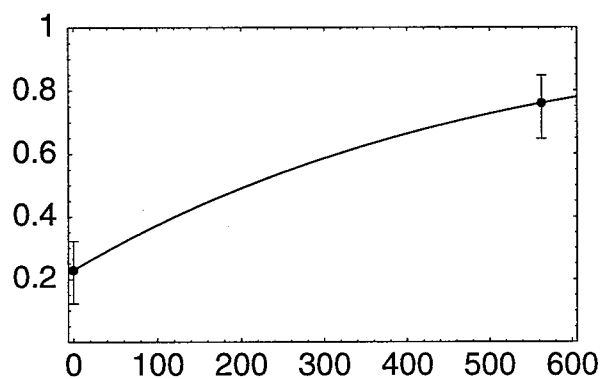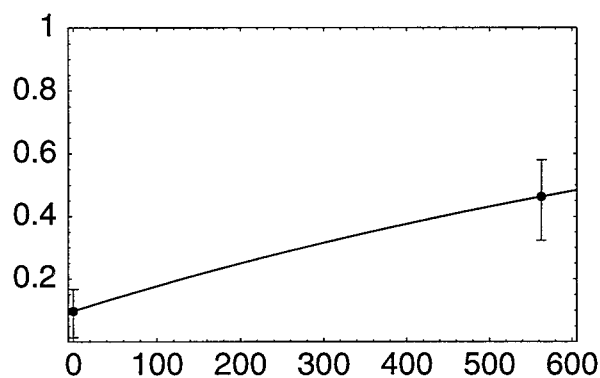The Optimized Function F of Dose d is:

F(d) = 1 - exp[-P(d)], where:

P(d) = 0.102654 + 0.000923402 d

ChiSquare(0) = 0. Perfect fit.

R^2 = 1.

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)

■ 5. NTP 1983 Rat F344/N: M 340 g   RenalTub Adenocarcinoma

```
xhi=300;
    doses = { 0, 198, 282}*1. (* mg/kg-d LTWAM *);
  ndosed = {33, 20, 16};
nrespond = { 0,  0,  3};
```

```
qTCE5=QFit[doses, ndosed, nrespond,  500,
   PolyDegree->2, Exponentiated->True,
   ConfInterval->.90, Output->Q1,
   Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->.5];
```

The Optimized Function F of Dose d is:
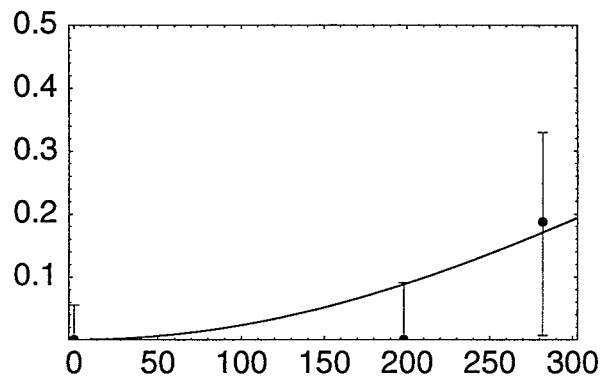
$F(d) = 1 - exp[-P(d)]$, where:

$P(d) = 2.35123 \times 10^{-6} d^2$

ChiSquare(2)= 1.96373   1-tailed p= 0.374612

$R^2 = 0.485332$

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)

■ 6. Bell et al. 1978 Mouse B6C3F1: M 35(?) g   HCC or HCA

```
xhi=300;
    doses = { 0, 42.3, 127, 254}*1. (* mg/kg-d LTWAM *);
   ndosed = {99, 95, 100, 97};
nrespond = {20, 35, 38, 53};
```

```
qTCE6=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->3, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

The Optimized Function F of Dose d is:

$F(d) = 1 - \exp[-P(d)]$, where:

$P(d) = 0.287436 + 0.00181511\,d + 2.65261 \times 10^{-9}\,d^3$

ChiSquare(1)= 3.29871 1-tailed p= 0.0693345

$R^2 = 0.881871$

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)
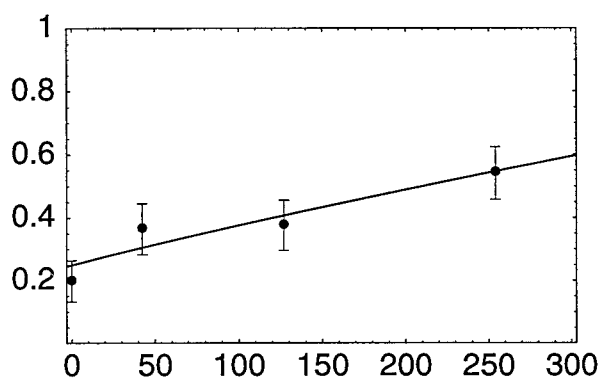
■ 7. Henchler et al. 1980 Mouse Han:NMRI: F 30(?) g   Malig. Lymphoma

Henchler did not consider this positive--called the study negative;
High spontaneous Malig. Lymphoma incidence is peculiar to this strain of mice in females (inborn murine lymphoma virus)
p=0.03 by Fisher Exact for females (this data set); p=1 for males

```
xhi=200;
    doses = { 0, 33.2, 166}*1. (* mg/kg-d LTWAM *);
   ndosed = {29, 30,    28};
 nrespond = { 9, 17,    18};
```

```
qTCE7=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->2, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->1];
```

The Optimized Function F of Dose d is:
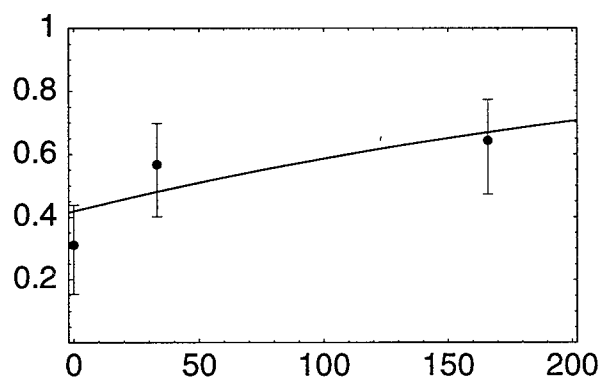
F(d) = 1 - exp[-P(d)], where:

P(d) = 0.54236 + 0.00338976 d

ChiSquare(1) = 2.3744   1-tailed p= 0.123339

R^2 = 0.673392

F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)

■ 8. Maltoni et al. 1986 Mouse Swiss: F 30(?) g   Malig. Hepatoma

```
xhi=250;
    doses = { 0, 35.3, 106, 212}*1. (* mg/kg-d LTWAM *);
  ndosed = {90, 90, 90, 90};
nrespond = { 4, 2, 8, 13};
```

```
qTCE8=QFit[doses, ndosed, nrespond,  500,
  PolyDegree->3, Exponentiated->True,
  ConfInterval->.90, Output->Q1,
  Xmin->-xhi/100,Xmax->1.01*xhi,Ymin->0,Ymax->.25];
```

The Optimized Function F of Dose d is:

F(d) = 1 - exp[-P(d)], where:

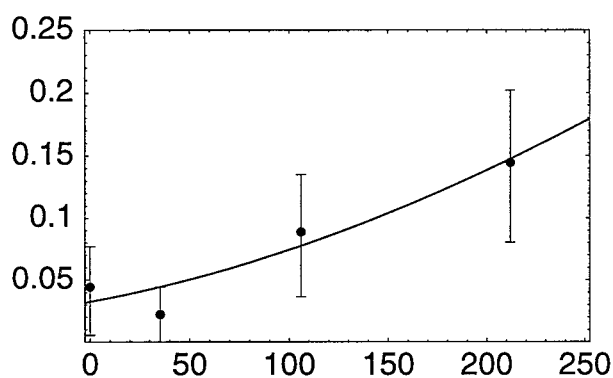$P(d) = 0.0329504 + 0.000306602\, d + 1.36839 \times 10^{-6}\, d^2$

ChiSquare(1)= 1.62806   1-tailed p= 0.201971

R^2 = 0.911347
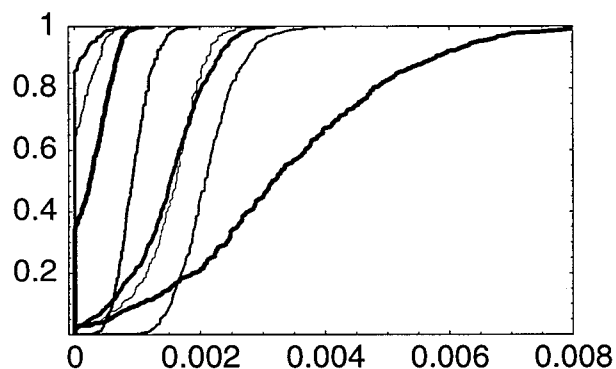
F(d) vs Data

(& Bootstrap 90% Conf. Limits on Data)


```
qall = {qTCE1, qTCE2, qTCE3, qTCE4, qTCE5, qTCE6,
qTCE7, qTCE8};
```

■ Weighted-Average TCE Cancer Potency

■ Define qall

```
cdfs = Cdf /@qall;
```

```
PlotCdf[cdfs, Xmin -> -.0001, Xmax -> .008];
```

Remove Henchler lymphoma data, due to his determination that this was a negative study given the likelihood of murine lymphoma virus involvement:

```
cdf7 = Drop[cdfs, {7}]; CdfQ/@cdf7
```

```
{True, True, True, True, True, True, True}
```

```
PlotCdf[cdf7, Xmin -> -.0001, Xmax -> .008];
```
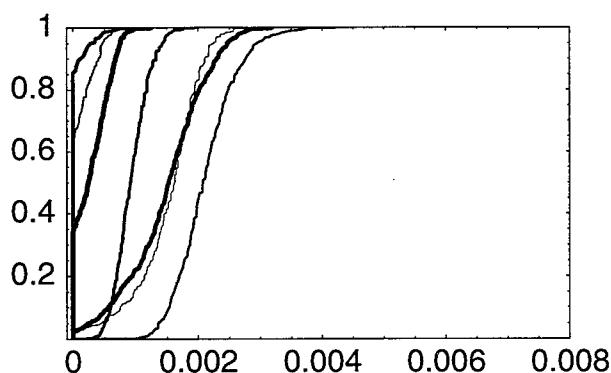

Replot to look nicer:

```
PlotCdf[cdf7, Xmin -> -10^-6, Xmax -> .00401];
```
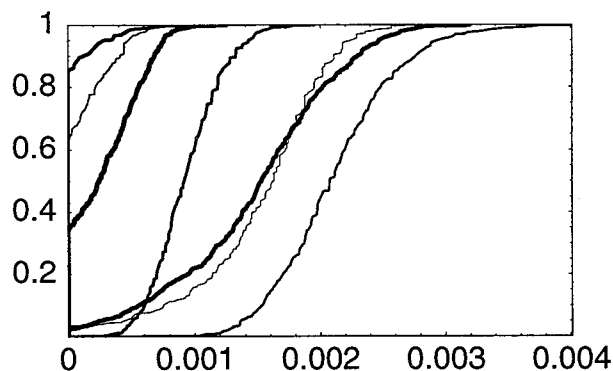


```
o = {EV[#], Edf[#, 0], Idf[#, {.5, .95}]}&/@cdf7;
```

```
oo = Prepend[Transpose[Prepend[Transpose[Flatten/@o],
    Range[7]]], {"Study", Mean, P0, "50th%ile", "95th%ile"}];
TBL[oo]
```

```
Study  Mean          P0      50th%ile      95th%ile
1      0.00151724    0.018   0.00162322    0.00224792
2      0.000100304   0.624   0             0.000474276
3      0.0021167     0       0.00207564    0.00294787
4      0.000933588   0       0.000915365   0.00142873
5      0.0000371087  0.854   0             0.000283281
6      0.00148119    0.028   0.00152863    0.00246978
7      0.000282383   0.342   0.000246207   0.000755992
```

Study-Weighting Logic:
{Species, Strain, Sex, Site, Study}-specific data are equally likely, therefore:
Data sets {{{{1,3,6},{2,4}}, 8}, 5}={mouse,rat}={mouse{b6c3f1{m,f},swiss},rat}
get relative weights: {1,1} = { { {{1,1,1}, {1.5,1.5}}, 6}, 12}

(Note that Henchler lymphoma data was removed, due to his determination that this was a negative study given the likelihood of murine lymphoma virus involvement.)

```
ww = { { {{1, 1, 1}, {3 / 2, 3 / 2}}, 6}, 12} / 24;
studies = {{{{1, 3, 6}, {2, 4}}, 8}, 5};
bwg = {{{{34, 37, 35}, {29, 33}}, 30}, 340};
wt = Transpose[Flatten /@ {studies, ww, bwg}];
swt = Sort[wt];
weights = #[[2]] & /@ swt;
bwg = Last /@ swt;
{ww, wt, swt, weights, bwg}
```

$$\{\{\{\{\{\frac{1}{24}, \frac{1}{24}, \frac{1}{24}\}, \{\frac{1}{16}, \frac{1}{16}\}\}, \frac{1}{4}\}, \frac{1}{2}\}, \{\{1, \frac{1}{24}, 34\}, \{3, \frac{1}{24}, 37\},$$
$$\{6, \frac{1}{24}, 35\}, \{2, \frac{1}{16}, 29\}, \{4, \frac{1}{16}, 33\}, \{8, \frac{1}{4}, 30\}, \{5, \frac{1}{2}, 340\}\}, \{\{1, \frac{1}{24}, 34\},$$
$$\{2, \frac{1}{16}, 29\}, \{3, \frac{1}{24}, 37\}, \{4, \frac{1}{16}, 33\}, \{5, \frac{1}{2}, 340\}, \{6, \frac{1}{24}, 35\}, \{8, \frac{1}{4}, 30\}\},$$
$$\{\frac{1}{24}, \frac{1}{16}, \frac{1}{24}, \frac{1}{16}, \frac{1}{2}, \frac{1}{24}, \frac{1}{4}\}, \{34, 29, 37, 33, 340, 35, 30\}\}$$

Average the study-specific cdfs using the study-specific weights ("weights") defined above, plot the results, and get statistics for the resulting averaged cdf ("adfBW") based on a body-weight (i.e., using a mass-per-kg-body-weight) approach to interspecies scaling of equitoxic doses.

```
adfBW = AverageCdf[cdf7, Weights -> weights];

adfBW >> "BogenHD:Ken:Projects:TCE Air Force:QbwCdf";

PlotCdf[adfBW, Xmin -> -10^-6, Xmax -> .003, Ymin -> .499];
```
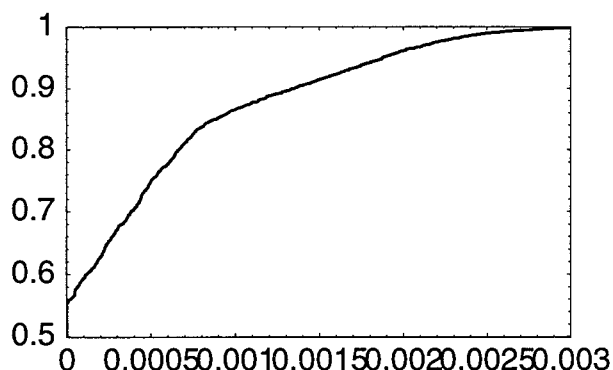


```
o1 = {Length[adfBW], EV[adfBW], Edf[adfBW, 0],
 Idf[adfBW, .95]};
TBL[{{Length, Mean, P0, "95th%ile"}, o1}]
```

Length	Mean	P0	95th%ile
1421.	0.000366899	0.553417	0.00187624

Multiply the abscissa of each cdf by $(Wh/BW)^{.25}$, where $Wh = 70$ kg and BW is rodent body weight in grams, i.e., scale using a $BW^{.75}$ scaling factor. Then re-average the cdfs using the same study-specific weights as used above, to obtain the resulting averaged cdf ("adf75") based on a (body weight)$^{.75}$ (i.e., using a mass-per-(kg body weight)$^{.75}$) approach to interspecies scaling of equitoxic doses.

```
bwr75 = (70 * 1000 / bwg) ^ .25
```

{6.73604, 7.00931, 6.59514, 6.7865, 3.78795, 6.6874, 6.95015}

```
cdf75 =
  Transpose /@ ({#[[1, 1]] * #[[2]], #[[1, 2]]} & /@ Transpose [{Transpose /@ cdf7, bwr75}]);
Dimensions /@ # & /@ {cdf7, cdf75}
```

{{{298, 2}, {62, 2}, {143, 2}, {109, 2}, {22, 2}, {488, 2}, {309, 2}},
 {{298, 2}, {62, 2}, {143, 2}, {109, 2}, {22, 2}, {488, 2}, {309, 2}}}

```
adf75 = AverageCdf[cdf75, Weights -> weights];
```

```
PlotCdf[{adf75, adfBW}, Xmin -> -10^-6, Xmax -> .02, Ymin -> .499];
```
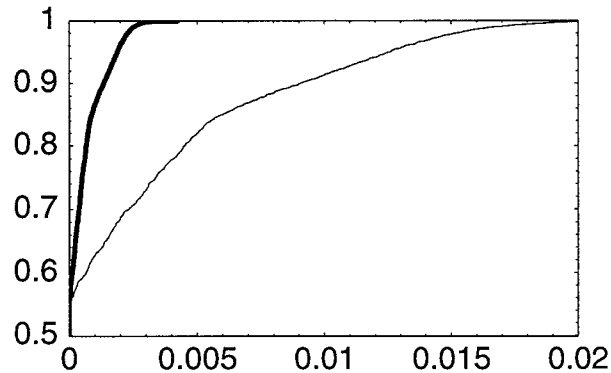


```
o2 = {Length[adf75], EV[adf75], Edf[adf75, 0],
 Idf[adf75, .95]};
TBL[{{Length, Mean, P0, "95th%ile"}, o2}]
```

```
Length  Mean        P0         95th%ile
1421    0.00242109  0.553417   0.0125937
```

Average the cdfs "adfBW" and "adf75" assuming equal likelihood, standardize, simplify, plot, get statistics, and save:

```
ade = AverageCdf[{adfBW, adf75}, Weights -> {.5, .5}];
ade = StandardizeCdf[ade, 500];
adf = SimplifyCdf[ade];
```

```
Dimensions /@ {ade, adf}
```

{{501, 2}, {226, 2}}

```
PlotCdf[{adf75, adfBW, adf}, Xmin -> -.00002, Xmax -> .02, Ymin -> .49];
```



```
o12 = {Length[adf], EV[adf], Edf[adf, 0], Idf[adf, .95]};
TBL[{{Length, Mean, P0, "95th%ile"}, o12}]
```

```
Length  Mean        P0     95th%ile
228     0.00140008  0.552  0.00902774
```
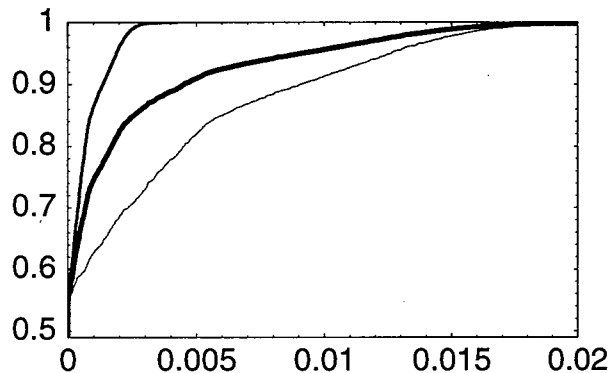
■ Define adf

```
WriteMatrix["BogenHD:Ken:Projects:TCE Air Force:Qcdf.txt", N[adf]];
```

TCE Threshold (Cytotoxicity) Model

■ TBARS dose-response in male B6C3F1 mice

27-g male B6C3F1 mice (Larson & Bull, 1992)
Dose in mg TCA per kg BW, vs. TBARS in nmol malondialdehyde equiv./g liver (n=4 @ ea. dose)

```
df = 3;
TBARScontrol = 40;
SDcontrol = 4;
dose = Log[10., {100, 300, 1000, 2000}];
tbars = {46, 67, 81, 108} - TBARScontrol;
```

$$sd = \sqrt{\{6, 7, 6, 7\}^2 + 4^2} \ ;$$

```
dat = Transpose[{dose, #}] & /@ {tbars + sd, tbars - sd, tbars};
{t95 = RQ[Q, T, 3, .95], TBARS95 = t95 * SDcontrol, TBARShi = Last[tbars]}
```

```
{2.35336, 9.41345, 68}
```

```
fxn[dose_, p_] := p[[1]] * NormalCdf[(dose - p[[2]]) / p[[3]]]
```

List "o" = {parameter estimates, corresponding SE values, $\{\chi^2, df, p\text{-value}\}$}:

```
o = LSMin[dose, tbars, {140, 4, 1}, fxn,
        Weights → sd^-2, KnownVariances → True, Progress → True, Step → 10^-4, Output → CVM]

0 {{140., 4., 1.}, 29.8243}

1 {{140., 3.996, 1.00361}, 29.2399}

2 {{140., 3.95962, 1.03601}, 24.2432}

3 {{140.001, 3.77426, 1.17963}, 7.75091}

4 {{140.004, 3.49797, 1.11006}, 2.23287}

5 {{140.007, 3.3745, 0.864436}, 1.33394}

6 {{140.023, 3.38641, 0.882103}, 1.3235}

7 {{140.175, 3.38757, 0.882599}, 1.32304}

{{140.023, 3.38641, 0.882103}, {{89321.2, 661.561, 283.881},
    {661.561, 4.91457, 2.1214}, {283.881, 2.1214, 0.945736}}, {1.32304, 1, 0.250047}}


vars = Diagonal[Sqrt[o[[2]]]];
vars = Table[vars[[i]] vars[[j]], {i, 3}, {j, 3}];
o[[2]] / vars

{{1., 0.998504, 0.976729}, {0.998504, 1., 0.983998}, {0.976729, 0.983998, 1.}}


Clear[fxn];
sdraw = √4 sd;
fxn[dose_, p_] := 100 * NormalCdf[(dose - p[[1]]) / p[[2]]]

o = LSMin[dose, tbars, {3, 1}, fxn, NYatX → {4, 4, 4, 4},
        SDY → sdraw, Weights → sdraw^-2,
    KnownVariances → True, Progress → True, Step → 10^-5, Output → CVM]

0 {{3., 1.}, 16.1584}

1 {{3.0001, 0.999935}, 16.1557}

2 {{3.00104, 0.999284}, 16.1294}

3 {{3.00964, 0.993005}, 15.8937}

4 {{3.05248, 0.944763}, 14.7945}

5 {{3.06417, 0.786705}, 13.6469}

6 {{3.05095, 0.731969}, 13.5613}

7 {{3.05227, 0.734029}, 13.5611}

{{3.05095, 0.731969},
    {{0.0084548, 0.00477488}, {0.00477488, 0.0311052}}, {13.5611, 14, 0.482896}}
```

Calculate the Pearson product-moment correlation of the two fitted lognormal-model parameters:

```
vars = Diagonal[Sqrt[o[[2]]]];
vars = Table[vars[[i]] vars[[j]], {i, 2}, {j, 2}];
o[[2]] / vars
```

{{1., 0.294438}, {0.294438, 1.}}

{0.00845479840778034485`, 0.0311052466621742773`}^.5

{0.09195, 0.176367}

```
o = LSMin[dose, tbars, {3, 1}, fxn, NYatX → {4, 4, 4, 4},
        SDY → sdraw, Weights → sdraw^-2,
   KnownVariances → True, Progress → True, Step → 10^-5]
```

0 {{3., 1.}, 16.1584}

1 {{3.0001, 0.999935}, 16.1557}

2 {{3.00104, 0.999284}, 16.1294}

3 {{3.00964, 0.993005}, 15.8937}

4 {{3.05248, 0.944763}, 14.7945}

5 {{3.06417, 0.786705}, 13.6469}

6 {{3.05095, 0.731969}, 13.5613}

7 {{3.05227, 0.734029}, 13.5611}

{{3.05095, 0.731969}, {0.09195, 0.176367}, {13.5611, 14, 0.482896}}

```
o = {{3.05095121221202791`, 0.731969108012972302`}, {0.0919499777475793855`,
      0.176366795804012603`}, {13.5610834321759421`, 14, 0.482895537121049489`}};
```

```
PlotData[dat, FitTo → {{fxn[d, o[[1]]], TBARS95 + 10^-6 * d}, d},
      Xmin → -.01, Xmax → 3.5, Ymin → -.01,
      Style → {M, J, M, OO}, DotSize → .015,
      Dashed → {False, .025}];
```
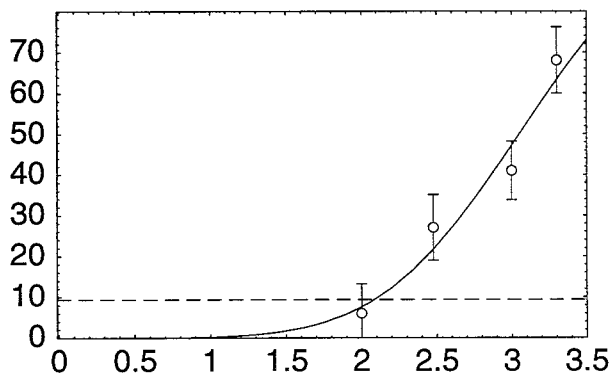

end

■ Effective acute TCA threshold dose (mg/kg) for TBARS elevation

■ **Define cdfD95 = Prob{significant TBARS elevation} I Effective TCA dose (mg/L plasma)**
(with abscissa units of mg TCA/L plasma based on Larson & Bull (TAP 115:268-277, 1992) using:
Vd = 15.0 mL
Cmax = 790 nmol TCA/mL plasma = 129.1 mg TCA/mL plasma
100 mg TCA in water administered by gavage)

To do this, first analytically solve for values of d95 I fxn[{p1,p2},d95] = TBARS95, where d95 is dose on a log10(mg/kg) scale, and fxn is the log-normal response function fitted above:
    100*NormalCdf[(dose-p1)/p2],
where p1 and p2 and their estimated standard deviations are repeated below, and where errors in p1 and p2 are assumed to have a joint T distribution with df = 14 a correlation r equal to ~0.294.

```
{{p1, p2} = o[[1]], {sdp1, sdp2} = o[[2]]}

{{3.05095, 0.731969}, {0.09195, 0.176367}}

{pval = TBARS95/ 100, zval = NormalCdf[TBARS95/ 100, Inv]}

{0.0941345, -1.31572}

d95 = Simplify[(p1 + sdp1*t14a) + (p2 + sdp2*t14b) *zval]

2.08789 + 0.09195 t14a - 0.232049 t14b

{constant= p1 +p2 * zval, w1 = sdp1, w2 = sdp2*zval}

{2.08789, 0.09195, -0.232049}
```

Therefore, d95 is distributed as a constant plus a weighted sum (v) of two correlated T-ditributed variates having weights w1 and w2 and degrees of freedom f1=14 and f2=14, respectively. A numerical-analytic solution for the cumulative distribution function $\psi(v)$ of the weighted sum (v) of two correlated T-variates is as follows, based on a straightforward extension of previously published results concerning independent weighted T-variates, to the case of correlated T-variates (Ruben, H. 1960. On the distribution of the weighted sum of two independent Student variables. *J. Royal Soc. Stat. Ser. B (Methodol.)* 22, 188-194.):

```
ψ[v_, f1_, f2_, w1_, w2_, ρ_: 0] := Module[{x, dBx, dIx, φx, f12 = f1 + f2},
    dBx = ∂ₓ Beta[x, f1/ 2, f2/ 2] / Beta[f1/ 2, f2/ 2];
    φx = Abs[Sqrt[
        (f12*x (1 - x)) / (f1 (1 - x) w1² + f2*x*w2² + 2 ρ*w1*w2 √(f1*f2*x (1 - x)) )]];
    If[NumberQ[v],
      NIntegrate[CDF[StudentTDistribution[f12], v*φx] *dBx,
        {x, 0, 1}, WorkingPrecision→ 20, AccuracyGoal→ 12, MaxRecursion→ 10],
      NIntegrate[CDF[StudentTDistribution[f12], #*φx] *dBx, {x, 0, 1},
        WorkingPrecision→ 20, AccuracyGoal→ 12, MaxRecursion→ 10] & /@ v
    ]
  ] /; NumberQ[v] || VectorQ[v, NumberQ]

{constant,w1,w2}={2.0878871554404643, 0.09194997774757939,
    -0.23204875723225765};
```

```
(*test*)
ψ[-1.5, 14, 14, w1, w2, 0.294]
```

4.869885×10^{-6}
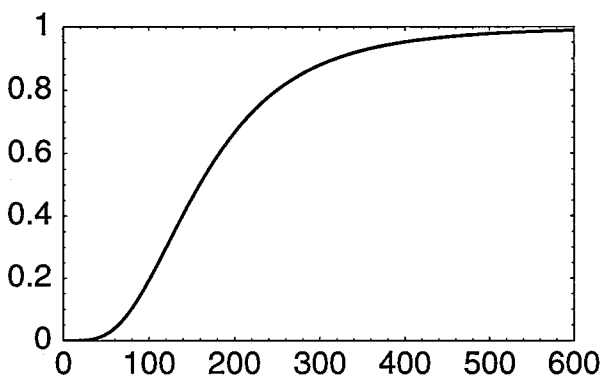
```
logdc = Join[-{5.5, 5, 4.5, 4, 3.5}, Range[-3, 0, .01]];
pp = ψ[#, 14, 14, w1, w2, 0.294] & /@ logdc;

pps = Join[pp, 1 - Rest[Reverse[pp]]];
logd = Join[logdc + constant, constant - Rest[Reverse[logdc]]];
cc = Transpose[{logd, pps}];
cc = Append[cc, {8, 1}];

cdfD95 = Prepend[{(10^#[[1]]) * (130./100), #[[2]]} & /@ cc, {0, 0}];

PlotCdf[cdfD95, Xmin→ -0.001, Xmax→ 600, Ymin→ -0.00001, Ymax→ 1];
```



```
cdfD95 >> "BogenHD:Ken:TCE Air Force:Data:TBARSvTCA";
```

■ **Obtain a low-risk analytic approximation of cdfD95 that will be used to calculate 2nd order approximations of cytotoxic risk (RcAng, RcBar), because cdfD95 is very nonlinear at low risks**

First look at cdfD95 in log-log space:

```
dat = Log[10., {#[[1]], #[[2]]}] & /@ Rest[cdfD95];

{First[dat], Last[dat]}

{{-3.29817, -12.5}, {8.11394, 0.}}

fdat = Take[dat, {4, 260}];
{First[fdat], Last[fdat]}

{{-1.79817, -10.5282}, {1.74183, -1.52916}}

Length[fdat]

257

wts = Join[{10^3, 10^3}, Table[1, {255}]];
```

```
fit = Fit[fdat, {1, x, x^2}, x]
```

$$-7.23848 + 2.46626\,x + 0.469885\,x^2$$

```
PlotData[dat, Xmin→ -2.01, Xmax→ 3, JoinPoints→ True,
    Ymin→ -10.01, Ymax→ -.00001, Style→ {OO}, DotSize→ .001, FitTo→ {fit, x}, Dashed→ .025];
```
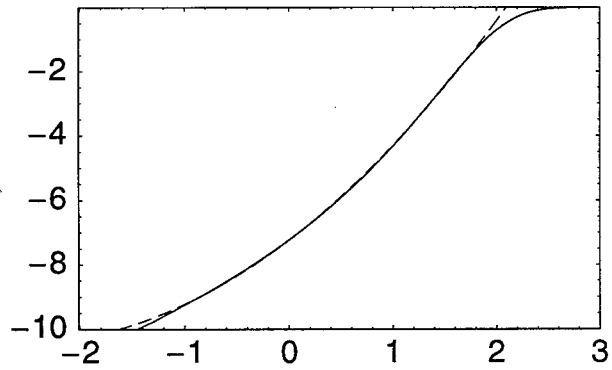

end

end

Appendix 3.H

TCE Risk

```
<< RiskQ`;

HardDrive = "BogenHD";
PathName[filename_, hardDrive_String: HardDrive] := Module[{file = filename},
        If[Head[file] =!= String, file = ToString[file]];
        StringJoin[hardDrive, ":Ken:TCE Air Force:Data:", file]
     ];
```

■ Log-Transform Utility Functions

```
GMGSDx::usage =
  "GMGSDx[Mx,SDx] returns the  geometric mean and geometric sdandard deviation of
    a lognormal variate X that also has the specified arithmetic mean Mx and
    arithmetic sdandard deviation SDx, based on the method of moments.";

MSDx::usage =
  "MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation
    of a lognormal variate X that also has the specified geometric mean GMx and
    geometric sdandard deviation GSDx, based on the method of moments.";

GMGSDx1::usage =
  "GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that
    the product X1*X2 has the desired coefficient of variation (CV) = cvWant,
    conditional on the lognormal variate X2 having an arithmetic mean and
    CV equal to 1 and cv2, respectively, based on the method of moments.";

MSDx[GMx_, GSDx_] := Module[{mux, sigy},
        sigy = Log[GSDx];
        mux = GMx E^((sigy^2) / 2);
        mux {1, Sqrt[E^(sigy^2) - 1]}]]

GMGSDx[Mx_, SDx_] := Module[{muy, sigy},
        sigy = Sqrt[Log[1 + (SDx / Mx)^2]];
        muy = Log[Mx] - (sigy^2) / 2;
        E^{muy, sigy}]
```

$$\text{GMGSDx1[cvWant\_, cv2\_] := Module}\Big[\{my1\},$$

$$muy1 = \text{Log}\Big[\text{Sqrt}\Big[\frac{cv2^2 + 1}{cvWant^2 + 1}\Big]\Big];$$

$$\text{E}^\wedge\{muy1, \text{Sqrt}[-2\,muy1]\}$$

$$\Big] \text{ /; cvWant >= cv2}$$

■ Extrapolation Factors (cf. Slob & Pieters, *Risk Anal.*, 1998; EPA)

Assumed that median is central target for uncertain EF, expected value is central target for heterogeneous EF.

EFinterspTdyn: Uncertain (Median =1)

     `{{z95, z99} = RQ[Q, N, {0, 1}, {.95, .99}], sdy = Log[3 / 1] / z99, gsd = E ^ sdy}`

     `{{1.64485, 2.32635}, 0.472248, 1.60359}`

     `{mux, sdx} = MSDx[1, gsd]; {mux, varx = sdx ^ 2}`

     `{1.11796, 0.312264}`

EFintraspTdyn: Heterogeneous (Mean = 1)

     `{Log[10 / 1] / z99, E ^ (Log[10 / 1] / z99)}`

     `{0.989785, 2.69066, 4.1787}`

     `rule = Solve[(mux / gmx) == E ^ (sigp * sigp / 2), gmx][[1]]`

     $\left\{ gmx \to E^{-\frac{sigp^2}{2}} \, mux \right\}$

     `Solve[sigp == (Log[xp / gmx] /. rule) / zp, sigp]`

     $\left\{ \left\{ sigp \to zp - \sqrt{zp^2 - 2\,Log\left[\frac{xp}{mux}\right]} \right\}, \left\{ sigp \to zp + \sqrt{zp^2 - 2\,Log\left[\frac{xp}{mux}\right]} \right\} \right\}$

     `10 ^ (2 / 3.)`

     `4.64159`

     $\left\{ sdy = z99 - \sqrt{z99^2 - 2\,Log[5]} \,,\; muy = -sdy * sdy / 2,\; gmx = E\verb|^|muy,\; gsdx = E \verb|^| sdy \right\}$

     `{0.845464, -0.357404, 0.69949, 2.32906}`

     `{gmx * E ^ (sdy * z99), 5 / gmx}`

     `{5., 7.14807}`

     `{mux, sdx} = MSDx[gmx, gsdx]; {mux, varx = sdx ^ 2}`

     `{1., 1.0438}`

(EF = 1+U): acute -> subchronic -> chronic EF, where EF>=1 and U is Uncertain (Median = 2x3 = 6)

     $\left\{ \{ sdyA = Log[6 / 3] / z99,\; gsd = E \verb|^| sdyA \}, \right.$

      $\left. \{ sdyB = Log[10 / 2] / z99,\; E \verb|^| sdyB \},\; \left\{ sdy = \sqrt{sdyA^2 + sdyB^2} \,,\; E \verb|^| sdy \right\} \right\}$

     `{{0.297955, 1.3471}, {0.69183, 1.99737}, {0.753264, 2.12392}}`

```
EFinterspTdyn = {LN, Log[{1, 1.60359450162908601`}]};
EFintraspTdyn = {LN, Log[{0.700, 2.33}]};
EFacuteTosubchr = {LN, Log[{3, 1.34710131239470021`}]};
EFsubchrTochr = {LN, Log[{2, 1.99736804456840957`}]};
EFacuteTochr1 = {LN, Log[{5, 2.12392092140740462`}]};

{mux, sdx} = MSDx[5, 2.123921]; {mux, varx = sdx^2}

{6.64019, 33.6726}


frisk[cdf_] := RQ[Cdf, cdf[[1]], cdf[[2]], 2000];
{cdfEFinter, cdfEFintra, cdfEFchron1} =
   frisk /@ {EFinterspTdyn, EFintraspTdyn, EFacuteTochr1};
{EFinterspTdynAng, EFintraspTdynBar, EFacuteTochr1Ang} =
  (EV[#, Empirical -> True] & /@ {cdfEFinter, cdfEFintra, cdfEFchron1})

{1.11699, 0.997245, 6.62175}


EFintraspTdynBar = 1;
```

Composite toxicodynamic uncertainty factor U = (UinterTdyn*(1 + Uacute2chr1)), has mean and variance:

```
mu1 = 1.11796373;
mu2 = 6.640194;
{muU = mu1 (1 + mu2), muU^2}

{8.54146, 72.9565}


v1 = 0.3122643770245;
v2 = 33.672622183427;
varU = (v1 * v2) + (mu1 * v2) + (mu2 * v1)

50.233
```

−see Ang and Tang, Probability Concepts in Statistics Vol. I, John Wiley & Sons, NY, p. 196.

```
end
```

Effective Dose

■ Genotoxic effective dose (EgBar, EgAng)

Exposures are all in units of (mg/kg-d) x 1000 (see "D. Effective Genotoxic Dose"):

```
Clear[cdfEingestBar, cdfEingestAng, cdfEdermalBar,
   cdfEdermalAng, cdfEinhaleBar, cdfEinhaleAng, EingestBar,
   EingestAng, EdermalBar, EdermalAng, EinhaleBar, EinhaleAng, o,
EingestBarAng, EinhaleBarAng, EdermalBarAng, EingestAngBar,
   EinhaleAngBar, EdermalAngBar];
```

```
o = Get /@ (PathName[#1] &) /@
     {EingestBar, EingestAng, EdermalBar, EdermalAng, EinhaleBar, EinhaleAng};
{cdfEingestBar, cdfEingestAng, cdfEdermalBar,
     cdfEdermalAng, cdfEinhaleBar, cdfEinhaleAng} = o;
{EingestBarAng, EingestAngBar, EdermalBarAng, EdermalAngBar,
   EinhaleBarAng, EinhaleAngBar} = (EV[#, Empirical → True] & /@ o);
o = {{EingestBarAng, EinhaleBarAng, EdermalBarAng},
   {EingestAngBar, EinhaleAngBar, EdermalAngBar}}
```

{{0.0507837, 0.00235084, 0.0071258}, {0.0491639, 0.00225019, 0.00684773}}
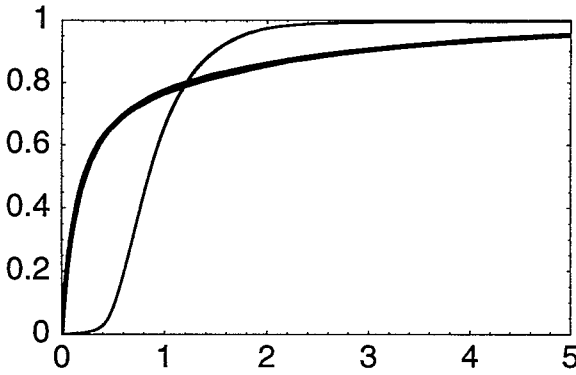
```
Plus @@ # & /@ o
```

{0.0602603, 0.0582618}

```
cdfBar = {({EingestBarAng⁻¹, 1} #) & /@ cdfEingestBar,
    ({EdermalBarAng⁻¹, 1} #) & /@ cdfEdermalBar, ({EinhaleBarAng⁻¹, 1} #) & /@ cdfEinhaleBar};
cdfAng = {({EingestAngBar⁻¹, 1} #) & /@ cdfEingestAng,
     ({EdermalAngBar⁻¹, 1} #) & /@ cdfEdermalAng, ({EinhaleAngBar⁻¹, 1} #) & /@ cdfEinhaleAng};
```

```
PlotCdf[Join[cdfBar, cdfAng], Xmin → -10⁻⁴, Xmax → 5, Ymin → -10⁻⁴];
```



```
Clear[EgBar, EgAng];
cdfBar = (Plus @@ cdfBar) / 3;
cdfAng = (Plus @@ cdfAng) / 3;
```

```
Put[cdfBar, PathName[EgBar]];
Put[cdfAng, PathName[EgAng]];
```

```
EgBar = cdfBar;
EgAng = cdfAng;
```

```
EgBar = Get[PathName[EgBar]];
EgAng = Get[PathName[EgAng]]; CdfQ /@ {EgBar, EgAng}
```

{True, True}

```
PlotCdf[{EgBar, EgAng}, Xmin → -10⁻⁴, Xmax → 5, Ymin → -10⁻⁴];
```



```
EV[#, Empirical → True] & /@ {EgBar, EgAng}
```

{1., 1.}

```
Var[#, Empirical → True] & /@ {EgBar, EgAng}
```

{2.45519, 4.42109}

```
{Edf[#, 1], Idf[#, {.5, .95, .99}]} & /@ {EgBar, EgAng}
```

{{0.658518, {0.844739, 1.74748, 2.59973}}, {0.769714, {0.223819, 4.81475, 10.326}}}

```
{{EingestAngBar, EinhaleAngBar, EdermalAngBar} /
   (Plus @@ {EingestAngBar, EinhaleAngBar, EdermalAngBar}),
 {EingestBarAng, EinhaleBarAng, EdermalBarAng} /
   (Plus @@ {EingestBarAng, EinhaleBarAng, EdermalBarAng})}
```
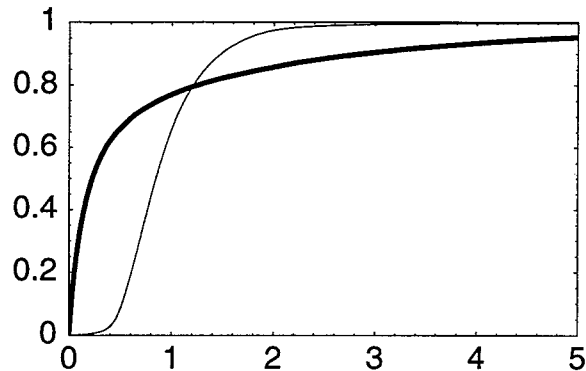
{{0.843844, 0.038622, 0.117534}, {0.842738, 0.0390114, 0.11825}}

Redefine {EingestAngBar,EdermalAngBar, EinhaleAngBar} each as a mean of the corresponding AngBar and BarAng means, then derive relative contributions of {EingestAngBar,EdermalAngBar, EinhaleAngBar} to Etotal, (where Etotal = EingestAngBar + EdermalAngBar + EinhaleAngBar).

```
{{EingestAngBar, EinhaleAngBar, EdermalAngBar} =
   ({EingestAngBar, EinhaleAngBar, EdermalAngBar} +
      {EingestBarAng, EinhaleBarAng, EdermalBarAng}) / 2,
     Etotal = Plus @@ {EingestAngBar, EinhaleAngBar, EdermalAngBar},
  FEingderinh = {EingestAngBar, EinhaleAngBar, EdermalAngBar} / Etotal}
```

{{0.0499738, 0.00230051, 0.00698677}, 0.0592611, {0.843282, 0.03882, 0.117898}}

```
Etotal = 0.0593 / 1000; (* mg/kg-d *)
{EingestAngBar, EinhaleAngBar, EdermalAngBar} = {0.843, 0.039, .118} * Etotal
```

{0.0000499899, 2.3127 × 10⁻⁶, 6.9974 × 10⁻⁶}

```
{EingestBarAng, EdermalBarAng, EinhaleBarAng} =
   {EingestAngBar, EdermalAngBar, EinhaleAngBar};
```

end

■ Cytotoxic exposures (EcBar, EcAng)

```
Clear[ECingestAng, ECingestBar, ECinhaleAng, ECinhaleBar, ECdermalAng, ECdermalBar];
ECingestAng = Get[PathName[ECingestAng]];
ECingestBar = Get[PathName[ECingestBar]];
ECinhaleAng = Get[PathName[ECinhaleAng]];
ECinhaleBar = Get[PathName[ECinhaleBar]];
ECdermalAng = Get[PathName[ECdermalAng]];
ECdermalBar = Get[PathName[ECdermalBar]];
CdfQ /@ {ECingestAng, ECingestBar, ECinhaleAng, ECinhaleBar, ECdermalAng, ECdermalBar}
```

{True, True, True, True, True, True}

```
{ECingestAngBar, ECingestBarAng, ECinhaleAngBar, ECinhaleBarAng,
  ECdermalAngBar, ECdermalBarAng} = EV[#, Empirical -> True] & /@
  {ECingestAng, ECingestBar, ECinhaleAng, ECinhaleBar, ECdermalAng, ECdermalBar}
```

{0.0161848, 0.0162664, 0.00835208, 0.00840379, 0.00226194, 0.0022786}

```
cdfCBar = {({ECingestBarAng⁻¹, 1} #) & /@ ECingestBar,
    ({ECdermalBarAng⁻¹, 1} #) & /@ ECdermalBar, ({ECinhaleBarAng⁻¹, 1} #) & /@ ECinhaleBar};
cdfCAng = {({ECingestAngBar⁻¹, 1} #) & /@ ECingestAng,
    ({ECdermalAngBar⁻¹, 1} #) & /@ ECdermalAng, ({ECinhaleAngBar⁻¹, 1} #) & /@ ECinhaleAng};
```

```
CdfQ /@ # & /@ {cdfCBar, cdfCAng}
```

{{True, True, True}, {True, True, True}}

```
PlotCdf[Join[cdfCBar, cdfCAng], Xmin → -10⁻⁴, Xmax → 5, Ymin → -10⁻⁴];
```
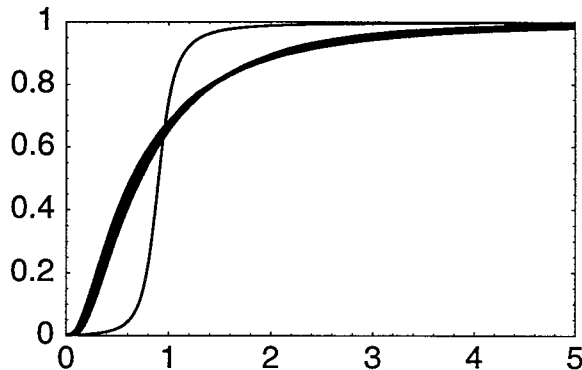


```
Dimensions /@ cdfCAng
```

{{405, 2}, {405, 2}, {405, 2}}

```
Clear[EcBar, EcAng];
cdfCBar = (Plus @@ cdfCBar) / 3;
cdfCAng = (Plus @@ cdfCAng) / 3;

Put[cdfCBar, PathName[EcBar]];
Put[cdfCAng, PathName[EcAng]];

EcBar = cdfCBar;
EcAng = cdfCAng;
```

```
EcBar = Get[PathName[EcBar]];
EcAng = Get[PathName[EcAng]]; CdfQ /@ {EcBar, EcAng}
```

{True, True}

```
PlotCdf[{EcBar, EcAng}, Xmin → -10⁻⁴, Xmax → 5, Ymin → -10⁻⁴];
```



```
EV[#, Empirical → True] & /@ {EcBar, EcAng}
```

{1., 1.}

```
Var[#, Empirical → True] & /@ {EcBar, EcAng}
```

{2.57639, 1.01933}

```
{Edf[#, 1], Idf[#, {.5, .95, .99}]} & /@ {EcBar, EcAng}
```
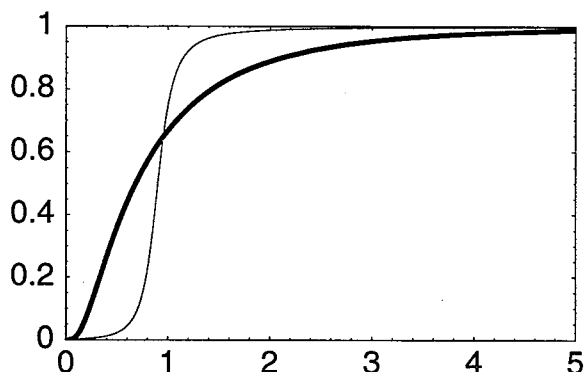
{{0.748303, {0.900642, 1.31192, 2.1726}}, {0.667287, {0.682963, 2.94235, 5.41263}}}

```
{{ECingestAngBar, ECinhaleAngBar, ECdermalAngBar} /
   (Plus @@ {ECingestAngBar, ECinhaleAngBar, ECdermalAngBar}),
 {ECingestBarAng, ECinhaleBarAng, ECdermalBarAng} /
   (Plus @@ {ECingestBarAng, ECinhaleBarAng, ECdermalBarAng})}
```

{{0.603938, 0.311658, 0.0844043}, {0.603605, 0.311842, 0.0845528}}

Define {CingestAngBar,CdermalAngBar, CinhaleAngBar} each as a mean of the corresponding AngBar and BarAng means, then derive relative contributions of {CingestAngBar,CdermalAngBar, CinhaleAngBar} to ECtotal, (where Ctotal = Cingest-AngBar + CdermalAngBar + CinhaleAngBar).

```
{{CingestAngBar, CinhaleAngBar, CdermalAngBar} =
   ({ECingestAngBar, ECinhaleAngBar, ECdermalAngBar} +
     {ECingestBarAng, ECinhaleBarAng, ECdermalBarAng}) / 2,
    Ctotal = Plus @@ {CingestAngBar, CinhaleAngBar, CdermalAngBar},
 FCingderinh = {CingestAngBar, CinhaleAngBar, CdermalAngBar} / Ctotal}
```

{{0.0162256, 0.00837794, 0.00227027}, 0.0268739, {0.603771, 0.31175, 0.0844788}}

```
Ctotal = 0.0269; (* mg TCA/L plasma *)
{CingestAngBar, CinhaleAngBar, CdermalAngBar} = {0.604, 0.312, .084} * Ctotal
```

{0.0162476, 0.0083928, 0.0022596}

```
{CingestBarAng, CdermalBarAng, CinhaleBarAng} =
  {CingestAngBar, CdermalAngBar, CinhaleAngBar};
```

end

end

Dose-Response

■ Genotoxic Potency (Qcdf)

```
Qcdf = Get[PathName[QbwCdf]];
QcdfAng = EV[Qcdf, Empirical -> True]
```

```
0.000366899
```

end

■ Cytotoxic Potency (TBARSvTCA)

```
TBARSvTCA = Get[PathName[TBARSvTCA]];
TBARSvTCAAng = EV[TBARSvTCA, Empirical → True]
```

```
186.389
```

```
{Idf[TBARSvTCA, {.5, .95, .99}], Edf[TBARSvTCA, 130]}
```

```
{{159.159, 395.291, 614.173}, 0.351186}
```

```
PlotCdf[TBARSvTCA, Xmin → -.0001, Xmax → 500, Ymin → -.001];
```
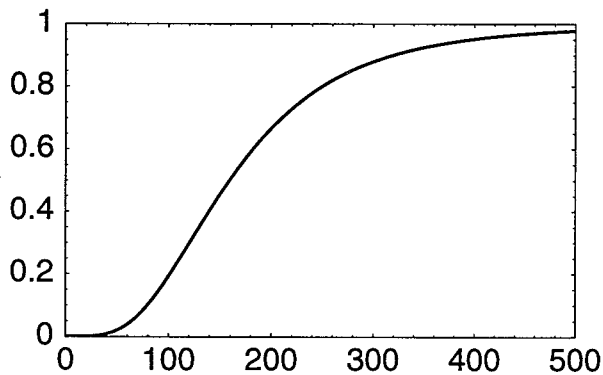


```
Last[TBARSvTCA]
```

```
{1.3 × 10⁸, 1}
```

$pTBARS[D_{Ca}] = F_C(D_{Ca})$ for effective acute cytotoxic dose D_{Ca}, i.e., the risk of significant TBARS elevation conditional on D_{Ca}.

```
pTBARS = Interpolation[Append[TBARSvTCA, {10⁹, 1}], InterpolationOrder → 1];
{pTBARS[130], pTBARS[0.0269]}
```

$$\{0.351186, 7.47274 \times 10^{-11}\}$$

end

end

TCE Risk (<R>, Rbar)

■ 2nd-order approximation terms for RBar and RAng

Let $\dfrac{10^a}{2} \to f$, $\text{Log}[10, x\,y\,z] \to L$, $\text{Log}[10, k\,x\,y\,z] \to Lk$,

and $\dfrac{2\,c}{\text{Log}[10]} \to \lambda$, where Log means ln (base e logarithm), and note that :

```
Log[x] / Log[10, x]
```

```
Log[10]
```

These terms are defined (see Bogen and Spear, 1987) as half the sum of second derivatives with respect to uncertain variates (for RAng), or to heterogeneous variates(for RBar), in the approximate risk function, which is linear-quadratic in $\text{Log}[10,-\text{Dca}]$, where Dca is acute effective cytotoxic dose. Note that, conditional on any variate x being evaluated at its expected value, the ratio $(\sigma^2)_x / x^2$ is equal to $(\gamma^2)_x$ where γ_x is the coefficient of variation of x (i.e., $\gamma_x = \sigma_x / \mu_x$). By using the latter relationship, simplified expressions derived below can be simplified further by incorporating corresponding coefficients of variation.

■ For RBar

```
d2 = (D[f * (k * x * y) ^ (b + c * Log[10, k * x * y]), {#, 2}] * (σ²)#) & /@ {x, y};
d2s = Simplify[Plus @@ d2]
```

$$\frac{1}{x^2\,y^2\,\text{Log}[10]^2}\left(f\,(k\,x\,y)^{b+\frac{c\,\text{Log}[k\,x\,y]}{\text{Log}[10]}}\right.$$
$$(\text{Log}[10]\,(2\,c + (-1+b)\,b\,\text{Log}[10]) + 2\,(-1+2\,b)\,c\,\text{Log}[10]\,\text{Log}[k\,x\,y] + 4\,c^2\,\text{Log}[k\,x\,y]^2)$$
$$\left.(y^2\,(\sigma^2)_x + x^2\,(\sigma^2)_y)\right)$$

in which k = DcaAngBar*Utdyn(1+Uchron)BcBar, x = V1 = Vtdyn, y = V2 = BcAng, and the result is evaluated conditional on V1 = V1Bar = VtdynBar = 1 and on V2 = V2Bar = BAngBar = 1. (Note that variates Ez are used to denote corresponding normalized dose variates Bz used elsewhere in this notebook, for all subscripts z.) In simplified form (along with a test that this form is correct):

$$\text{d2Simplified} = \frac{10^a}{2}\,(k\,x\,y)^{b+c\,L}\,(\lambda + b\,(b-1) + 2\,c\,L\,(2\,(b+c\,L)-1))\left(\frac{(\sigma^2)_x}{x^2} + \frac{(\sigma^2)_y}{y^2}\right);$$

```
{x1, x2} = N[{d2s, d2Simplified}] //. {a → -2., b → .05, c → .15, x → .2, y → .1,
```

$$(\sigma^2)_x \to 1.1, \ (\sigma^2)_y \to 1.8, \ L \to Log[10, k\, x\, y], \ \lambda \to \frac{2\, c}{Log[10]}, \ k \to .069, \ f \to \frac{10^a}{2}\}$$

```
{20.0338, 20.0338}
```

■ For RAng

```
d4 = (D[f * (k * u (x + y + z)) ^ (b + c * Log[10, k * u (x + y + z)]), {#, 2}] * (σ²)#) & /@ {u, x, y, z};
d4s = Simplify[Plus @@ d4]
```

$$\left(f\, (k\, u\, (x + y + z))^{b + \frac{c\, Log[k\, u\, (x+y+z)]}{Log[10]}} \ (Log[10]\, (2\, c + (-1 + b)\, b\, Log[10]) + \right.$$
$$2\, (-1 + 2\, b)\, c\, Log[10]\, Log[k\, u\, (x + y + z)] + 4\, c^2\, Log[k\, u\, (x + y + z)]^2)$$
$$\left.((x + y + z)^2\, (\sigma^2)_u + u^2\, ((\sigma^2)_x + (\sigma^2)_y + (\sigma^2)_z))\right) \Big/ (u^2\, (x + y + z)^2\, Log[10]^2)$$

in which k = DcaAngBar*Vtdyn*BcAng, u = U1 = Utdyn(1+Uchron), x = U2 = fCing*BcBar, y = U3 = fCinh*BcBar, z = U4 = fCder*BcBar, and the result is evaluated conditional on U1 = U1Bar = [Utdyn(1+Uchron)]Ang and on Ui = UiAng = fCi*BcBarAng = fCi for i = {ing, inh, der}. (Note that variates Ez are used to denote corresponding normalized dose variates Bz used elsewhere in this notebook, for all subscripts z.) In simplified form (along with a test that this form is correct):

```
d4Simplified =
```
$$\frac{10^a}{2}\, (k * u\, (x + y + z))^{b + c\, L}\, (\lambda + b\, (b - 1) + 2\, c\, L\, (2\, (b + c\, L) - 1)) \left(\frac{(\sigma^2)_u}{u^2} + \frac{(\sigma^2)_x + (\sigma^2)_y + (\sigma^2)_z}{(x + y + z)^2}\right);$$

```
{x1, x2} = N[{d4s, d4Simplified}] //.
   {a → 2., b → .5, c → .1, x → .2, y → .3, z → .4, u → .04, (σ²)u → 0.05, (σ²)x → 1.1,
```
$$(\sigma^2)_y \to 1.4, \ (\sigma^2)_z \to 1.7, \ L \to Log[10, k*u\, (x + y + z)], \ \lambda \to \frac{2\, c}{Log[10]}, \ k \to .069, \ f \to \frac{10^a}{2}\}$$

```
{46.8912, 46.8912}
```

Estimates of variance (e.g., in Bc) are used to evaluate the above expressions; e.g., recall that:

```
Var[#, Empirical → True] & /@ {EcBar, EcAng}

{2.57639, 1.01933}
```

end

■ Rang (Variability Distribution)

{EFintraspTdyn, EgAng x 3, EcAng x 3} = 7 heterogeneous variates
Ug = Likelihood that Rg is true ~ U[0, 0.5] by assumption; therefore, UgAng = 1/4.
{a,b,c,} = estimates of linear-quadratic parameters in Log[10, Dca] that model Fc(Dca) (see "Potency.nb")

Dose rank correlations (from "F. Effective Dose Correlations") **for:**   **(n = 50 x 500)**

GingAng	**CingAng**	**GinhAng**	**CinhAng**	**GderAng**	**CderAng**
1.	0.233	0.878	-0.00269	0.894	0.000496
0.233	1.	-0.0088	0.421	0.00439	0.514
0.878	-0.0088	1.	0.187	0.918	0.035
-0.00269	0.421	0.187	1.	0.0766	0.649
0.894	0.00439	0.918	0.0766	1.	0.177
0.000496	0.514	0.035	0.649	0.177	1.

Assign corresponding values to the upper triangular portion of the simulation-input-variate rank-correlation matrix (which shall be denoted "corr"). Note that the first row of the matrix pertains to the EFintraspTdyn variate (i.e., the intraspecies toxicodynamic extrapolation factor), which is not correlated with any of the 6 exposure variates.

```
corr = {{0, 0, 0, 0, 0, 0}, {.23, .88, 0, .89, 0},
    {0, .42, 0, .51}, {0.19, .92, 0.035}, {.077, .65}, {.18}};
```

Verify that the Cholesky decomposition of the target rank-corrlation matrix (=Reflect[corr]) contains no imaginary parts:

```
Cholesky[Reflect[corr]]

{{1, 0, 0, 0, 0, 0, 0}, {0, 1, 0, 0, 0, 0, 0}, {0, 0.23, 0.973191, 0, 0, 0, 0},
    {0, 0.88, -0.207976, 0.42702, 0, 0, 0}, {0, 0, 0.43157, 0.655136, 0.620116, 0, 0},
    {0, 0.89, -0.210339, 0.217916, 0.0403333, 0.338442, 0},
    {0, 0, 0.524049, 0.337196, 0.32724, 0.601429, 0.37798}}
```

```
Ux = 8.54146;
```

$$\text{sumsig} = \left(\frac{50.233}{\text{Ux}^2} + (0.604^2 + 0.312^2 + 0.084^2) * 2.5764 \right)$$

1.89742

```mathematica
{nsam, nsim} = {2000, 10};
UgAng = 1 / 4; (* see Rbar section below *)
cdfs = {EFintraspTdyn, EgAng, EcAng, EgAng, EcAng, EgAng, EcAng};
Clear[fxn];
fxn[x_, Ging_, Cing_, Ginh_, Cinh_, Gder_, Cder_, ratio_: False] := Module[
    {Rg, Rc1, Rc2, Rc, Rtot, k = 0.0269, Call, λ, V, L,
    a = -7.23848, b = 2.46626, c = 0.469885, Ux = 8.54146, sumsig = 1.89742},
 Rg = 1 - E ^ - (QcdfAng * EFinterspTdynAng * x * 0.0000593 *
            (0.843 Ging + 0.039 Ginh + 0.118 Gder));
    Ux = 8.54146;
    Call = 0.604 Cing + 0.312 Cinh + 0.084 Cder;
    V = x * Ux * k * Call;
    Rc1 = pTBARS[V];
    λ = 2 c / Log[10];
    L = Log[10, V];
    Rc2 = 10^a/2 V^(b+c*L) (λ + b (b - 1) + 2 c L (2 (b + c L) - 1)) sumsig;
    Rc = Rc1 + Rc2;
    Rtot = 1 - (1 - UgAng * Rg) (1 - Rc);
    If[ratio, UgAng * Rg / Rtot, Rtot]
    ];

fxnFG[x_, Ging_, Cing_, Ginh_, Cinh_, Gder_, Cder_] :=
 fxn[x, Ging, Cing, Ginh, Cinh, Gder, Cder, True]


EFintraspTdynBar = 1; {EFintraspTdynBar, EingestAngBar,
 EinhaleAngBar, EdermalAngBar, CingestAngBar, CinhaleAngBar, CdermalAngBar}
```

$\{1, 0.0000499899, 2.3127 \times 10^{-6}, 6.9974 \times 10^{-6}, 0.0162476, 0.0083928, 0.0022596\}$

```mathematica
fxn[EFintraspTdynBar, 1, 1, 1, 1, 1, 1]
```

1.31708×10^{-8}

```mathematica
fxnFG[EFintraspTdynBar, 1, 1, 1, 1, 1, 1]
```

0.461294

```mathematica
Timing[sim1 = Table[SimulateCdf[cdfs, nsam, Correlate → corr, Report → Append], {nsim}];]
```

{117.017 Second, Null}

```mathematica
(* Check Jennrich X2 p-values *)
jenp = Transpose[{Range[Length[sim1]], Last /@ Last /@ Last /@ sim1}]
```

{{1, 0.998878}, {2, 0.999285}, {3, 0.971796}, {4, 0.99744}, {5, 0.986883},
 {6, 0.99969}, {7, 0.998695}, {8, 0.986122}, {9, 0.998931}, {10, 0.827001}}

```mathematica
{jen, cdf, cvm} =
  QUAnalyze[cdfs, fxn, nsam, nsim, SimIn → sim1, Correlate → corr, Scale → 10^6];
```

TBL /@ jen

					Fractile	Value	CVM(%)
					0.01	0.0000790371	2.52229
					0.05	0.000231529	2.6391
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj	0.5	0.00460686	1.47592
0.000176653	0.0187711	14.9189	21	0.99969	0.95	0.172878	2.78787
					0.99	1.36419	8.07152
					Mean	0.230061	43.1698
					Variance	138.397	94.2987

PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → 1];
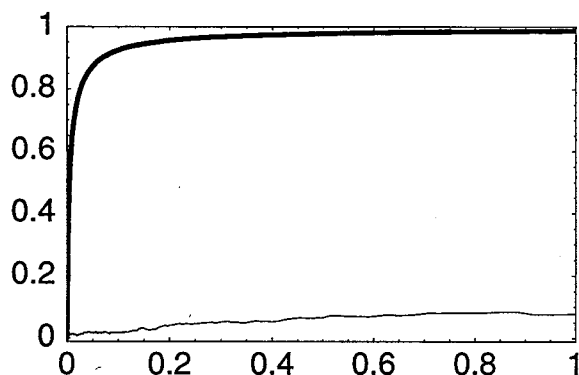
Put[cdf, PathName[Rang]]; Rang = cdf;

Rang = Get[PathName[Rang]]; CdfQ[Rang]

True

{RangBar = EV[Rang, Empirical → True],
{Rang50, Rang95, Rang99} = Idf[Rang, {.5, .95, .99}], Edf[Rang, 1]}

{0.230061, {0.00460686, 0.172878, 1.36419}, 0.987149}

Fraction of total risk due to genotoxic risk:

{jen, cdf, cvm} =
QUAnalyze[cdfs, fxnFG, nsam, nsim, SimIn → sim1, Correlate → corr, Scale → 1];

TBL /@ jen

					Fractile	Value	CVM(%)
					0.01	0.00262033	5.59528
					0.05	0.0182746	1.61415
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj	0.5	0.465644	0.531193
0.000466053	0.0212522	16.3988	21	1.	0.95	0.967092	0.0530831
					0.99	0.99007	0.0407695
					Mean	0.479793	0.127858
					Variance	0.107487	0.320455

end

■ Rbar (Uncertainty Distribution)

{Ug, EgBar, EcBar, Qcdf, EFinterspTdyn, EFacuteTochr1} = 6 uncertain variates

Ug = Likelihood that Rg is true ~ U[0, 0.5] by assumption

{a,b,c,} = estimates of linear-quadratic parameters in Log[10, Dca] that model Fc(Dca) (see "Potency.nb")

```
corr = {{0, 0, 0, 0, 0}, {.49, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
```

$$\text{sumsig} = \left(\frac{1.0438}{1^2} + \frac{1.01933}{1^2}\right)$$

```
2.06313
```

```
{nsam, nsim} = {2000, 10};
Ug = {{0, 0}, {.5, 1}};
cdfs = {Ug, EgBar, EcBar, Qcdf, EFinterspTdyn, EFacuteTochr1};
Clear[fxn];
fxn[ug_, Gall_, Call_, q_, uf1_, uf2_, ratio_: False] := Module[
        {Rg, Rc1, Rc2, Rc, Rtot, k = 0.0269,
         λ, U, L, a = -7.23848, b = 2.46626, c = 0.469885, sumsig = 2.06313},
        Rg = 1 - E^-(q*uf1*1*0.0000593 Gall);
      U = k*uf1*(1 + uf2)*Call;
      Rc1 = pTBARS[U];
      λ = 2 c / Log[10];
      L = Log[10, U];
```

$$\text{Rc2} = \frac{10^a}{2} \, U^{b+c*L} \, (\lambda + b \, (b-1) + 2\,c\,L \, (2 \, (b+c\,L) - 1)) * \text{sumsig};$$

```
      Rc = Rc1 + Rc2;
       Rtot = 1 - (1 - ug*Rg) (1 - Rc);
       If[ratio, ug*Rg / Rtot, Rtot]
          ];
```

```
fxnFG[ug_, Gall_, Call_, q_, uf1_, uf2_] := fxn[ug, Gall, Call, q, uf1, uf2, True]
```

```
{QcdfAng, EFinterspTdynAng, EFacuteTochr1Ang}
```

```
{0.000366899, 1.11699, 6.62175}
```

```
fxn[0.25, 1, 1, QcdfAng, EFinterspTdynAng, EFacuteTochr1Ang]
```

$$1.3516 \times 10^{-8}$$

```
fxnFG[0.25, 1, 1, QcdfAng, EFinterspTdynAng, EFacuteTochr1Ang]
```

```
0.449514
```

```
Timing[{jen, cdf, cvm} = QUAnalyze[cdfs, fxn, nsam, nsim, Correlate → corr, Scale → 10^6];]
```
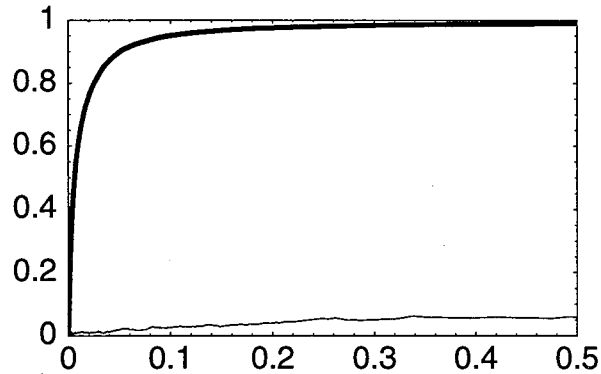
```
{79.5833 Second, Null}
```

TBL /@ jen

					Fractile	Value	CVM(%)
					0.01	0.000123904	3.00105
					0.05	0.000341113	1.72923
Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj	0.5	0.00608081	0.948997
7.27327×10^{-6}	0.0189821	2.71919	15	1.	0.95	0.0983222	2.46136
					0.99	0.491497	5.9934
					Mean	0.915155	68.8445
					Variance	5667.36	97.831

PlotCdf[{cvm, cdf}, Ymin $\rightarrow$ -.0001, Xmin $\rightarrow$ -.0001, Xmax $\rightarrow$ .5];

Clear[Rbar];
Put[cdf, PathName[Rbar]]; Rbar = cdf;

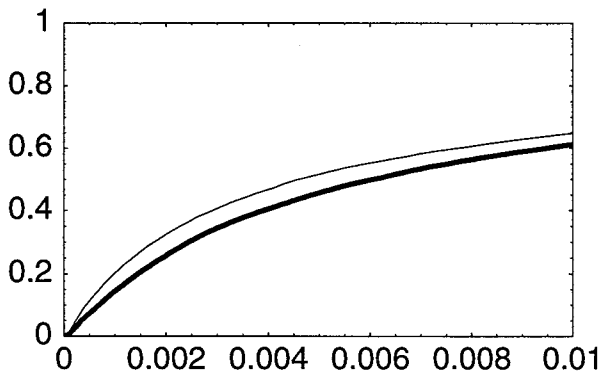**Rbar = Get[PathName[Rbar]]; CdfQ[Rbar]**

True

{RbarAng = EV[Rbar, Empirical $\rightarrow$ True],
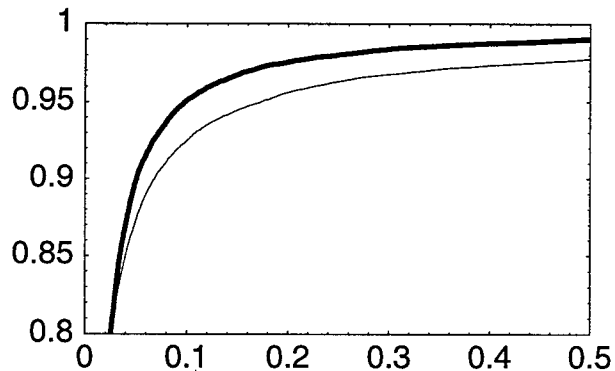 **{Rbar50, Rbar95, Rbar99} = Idf[Rbar, {.5, .95, .99}], Edf[Rbar, 1]}**

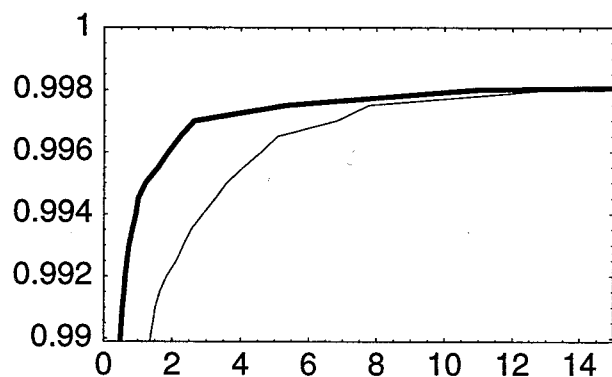{0.915155, {0.00608081, 0.0983222, 0.491497}, 0.994501}

end

PlotCdf[{Rang, Rbar}, Ymin $\rightarrow$ -.0001, Xmin $\rightarrow$ -.000001, Xmax $\rightarrow$ .01];

`PlotCdf[{Rang, Rbar}, Ymin → .7999, Xmin → -.000001, Xmax → .5];`

`PlotCdf[{Rang, Rbar}, Ymin → .9899, Xmin → -.000001, Xmax → 15];`

Fraction of total risk due to genotoxic risk:

`{jen, cdf, cvm} = QUAnalyze[cdfs, fxnFG, nsam, nsim, Correlate → corr, Scale → 1];`

`TBL /@ jen`

Mean[Δr]	Max[|Δr|]	JennrichChi2	DegFr	P-adj
0.000213108	0.0170892	2.45815	15	0.999999

Fractile	Value	CVM(%)
0.01	0	0
0.05	0	0
0.5	0	0
0.95	0.946036	0.125971
0.99	0.984215	0.0656233
Mean	0.253274	0.356907
Variance	0.126086	0.343434

`Edf[cdf, 0]`

0.553223

end

Confidence Bounds on JUV in Risk

- **R*99 = Analytic upper-bound JUV estimator (@ 99th %ile on U & V)**

      {Rang99, RangBar, Rbar99, RbarAng}

      {1.36419, 0.230061, 0.491497, 0.915155}

R*99b = rho99 x Rang99, rho99 = (Rbar99)/(\<Rbar\>)   \<------ As defined by Bogen (1995)
R*99a = rho99 x Rbar99, rho99 = (\<R\>99)/(\<R\>bar)   \<------ Alternative definition

      (* CV for R99* *)
      {RbarAngcv, Rbar99cv, RangBarcv, Rang99cv} = {.688, .0599, .432, .0807};

      {rhoa = (Rang99 / RangBar), rhob = (Rbar99 / RbarAng)}

      {5.92966, 0.537064}

$$\left(\left(1 - (1 - \#[[1]]^2)\,(1 - \#[[2]]^2)\right)^{.5}\right) \,\&\, /@\, \{$$
              {Rang99cv, RangBarcv}, {Rbar99cv, RbarAngcv}}

      {0.438088, 0.689372}

      {R99a = Rbar99 * rhoa, R99b = Rang99 * rhob, R99a / R99b}

      {2.91441, 0.732656, 3.97787}

$$\{\text{gammaa, gammab}\} = \left(\left(\left(1 - (1-\#[[1]]^2)\,(1-\#[[2]]^2)\,(1-\#[[3]]^2)\right)^{.5}\right) \,\&\, /@\, \{$$
              {Rbar99cv, Rang99cv, RangBarcv}, {Rang99cv, Rbar99cv, RbarAngcv}})

      {0.441385, 0.691846}

- **R*99 = Target-Nested Monte-Carlo JUV estimators (@ 99th %ile on V)**

```
SimulateCdfs[Cdfs_, nsam_, nsim_, options___] := Module[{o, cdfs = Cdfs, x},
        o = SimulateCdf[cdfs, nsam, options];
        If[Head[o] === String, Return[StringJoin["SimulateCdfs: Bad input\n", o]]];
        x = If[Dimensions[o] === {Length[cdfs], nsam}, o, o[[1]]];
        cdfs = (Cdf[#1, Xmax → 10^15] &) /@ x;
    Prepend[Table[SimulateCdf[cdfs, nsam, options], {nsim - 1}], o]
    ] /; nsam > nsim > 1
```

```
corrV = {{0, 0, 0, 0, 0, 0}, {.23, .88, 0, .89, 0}, {0, .42, 0, .51}, {0.19, .92,
   0.035}, {.077, .65}, {.18}};
cdfV = {EFintraspTdyn, EgAng, EcAng, EgAng, EcAng, EgAng, EcAng};
fxnV[x_, Ging_, Cing_, Ginh_, Cinh_, Gder_, Cder_] := Module[
        {Rg, Rc},
        Rg =
     1 - E ^ - (QcdfAng * EFinterspTdynAng * x * 0.0000593 (0.843 Ging + 0.039 Ginh + 0.118 Gder));
        Rc = pTBARS[EFinterspTdynAng * (1 + EFacuteTochr1Ang
                    ) * x * 0.0269 (0.604 Cing + 0.312 Cinh + 0.084 Cder)];
        1 - (1 - 0.25 Rg) (1 - Rc)
        ];

corrU = {{0, 0, 0, 0, 0}, {.49, 0, 0, 0}, {0, 0, 0}, {0, 0}, {0}};
cdfU = {Ug, EgBar, EcBar, Qcdf, EFinterspTdyn, EFacuteTochr1};
fxnUV[ug_, Gall_, Call_, q_, uf1_, uf2_,
        x_, Ging_, Cing_, Ginh_, Cinh_, Gder_, Cder_] := Module[
        {Rg, Rc},
     Rg = 1 - E ^ - (q * uf1 * x * 0.0000593 Gall (0.843 Ging + 0.039 Ginh + 0.118 Gder));
        Rc = pTBARS[uf1 * (1 + uf2) * x *
                    0.0269 Call (0.604 Cing + 0.312 Cinh + 0.084 Cder)];
        1 - (1 - ug * Rg) (1 - Rc)
        ];

fxnUV[0.25, 1, 1, QcdfAng, EFinterspTdynAng,
 EFacuteTochr1Ang, EFintraspTdynBar, 1, 1, 1, 1, 1, 1]
```

8.5699×10^{-9}

```
{nsam, nsim} = {999, 100}; {{i50, i95, i99} = (nsam + 1) {50, 95, 99} / 100}
```

```
{{500, 950, 990}}
```

```
test = Cdf[2 Range[nsam], Xmax -> 10 * nsam];
{Idf[test, .99], test[[i99 + 1]]}
```

$\left\{1980, \left\{1980, \frac{99}{100}\right\}\right\}$

```
(CdfQ[#] || RQ[Test, #[[1]], #[[2]]]) & /@ cdfV
```

```
{True, True, True, True, True, True, True}
```

```
(CdfQ[#] || RQ[Test, #[[1]], #[[2]]]) & /@ cdfU
```

```
{True, True, True, True, True, True}
```

```
Timing[
simv = SimulateCdfs[
        cdfV, nsam, nsim, Correlate -> corrV, Report -> False];
rvi = fxnV @@ # & /@ simv;
vi = (#[[i99, 2]] & /@ (Sort /@ MapThread[
        Transpose[{#1, Transpose[#2]}] &, {rvi, simv}]));
simu = SimulateCdfs[cdfU, nsam, nsim, Correlate -> corrU, Report -> False];
vij = Transpose[Table[#, {nsam}]] & /@ vi;
simin = MapThread[Join[#1, #2] &, {simu, vij}];
Dimensions[simin]]
```

```
{499.067 Second, {100, 13, 999}}
```

```
Timing[{jen, cdf, cvm} = QUAnalyze[13, fxnUV, nsam, nsim, SimIn → simin, Scale → 10^6];]
```
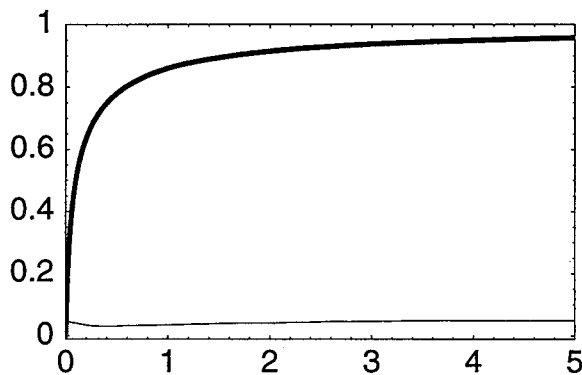
{35.0667 Second, Null}

```
min (499.067 + 35.067) / 60
```

8.90223 min

```
TBL[jen]
```

```
Fractile  Value          CVM(%)
0.01      0.00191737     4.40108
0.05      0.00557622     4.31837
0.5       0.0990887      4.15774
0.95      4.08454        4.93046
0.99      32.9278        5.67837
Mean      121.391        17.1954
Variance  5.42842×10^7   27.8288
```

```
PlotCdf[{cvm, cdf}, Ymin → -.01, Xmin → -.0001, Xmax → 5];
```


end

■ R*99 = Traditional Nested Monte-Carlo JUV estimators (@ 99th %ile on V)

Compare variability fractiles for the 50th, 95th and 99th %ile with respect to uncertainty, respectively, obtained using a traditional nested Monte-Carlo approach:

```
{Dimensions /@ {simu, simv}, {nsam, nsim, i50, i95, i99}}
```

{{{100, 6, 999}, {100, 7, 999}}, {999, 100, 500, 950, 990}}

```
Timing[o = Transpose /@ Table[
              Prn[i]; Table[
uij = Transpose[Table[#[[j]] & /@ simu[[i]], {nsam}]];
simin = Join[uij, simv[[i]]];
ruj = Sort[fxnUV @@ simin][[{i50, i95, i99}]],
{j, nsam}], {i, nsim}];]
```

i= 100

{34278. Second, Null}

```
{38278. (60.^-2) h, 38278. / (499.067 + 35.067)}
```

```
{10.6328 h, 71.6637}
```

```
{Dimensions[o], nsam, nsim}
```

```
{{100, 3, 999}, 999, 100}
```

```
fx[r_] := r
```

```
in = List /@ # & /@ Transpose[o];
Dimensions[in]
```

```
{3, 100, 1, 999}
```

```
{jen, cdf, cvm} = Transpose[QUAnalyze[1, fx, nsam, nsim, SimIn → #, Scale → 10^6] & /@ in];
```

Variability fractiles for the 50th, 95th and 99th %ile with respect to uncertainty, respectively:

```
TBL /@ jen
```

```
  Fractile  Value          CVM(%)
  0.01      0.0000280892   1.40513
  0.05      0.0000751209   0.711428
 ⌠0.5       0.000966502    0.385091
 ⌡0.95      0.0117496      0.721707 '
  0.99      0.033135       1.59806
  Mean      0.0292517      34.0115
  Variance  10.4468        69.2366
```

```
  Fractile  Value          CVM(%)     Fractile  Value          CVM(%)
  0.01      0.000642227    1.55942    0.01      0.00223991     2.71725
  0.05      0.00150738     1.19946    0.05      0.006994       2.28426
  0.5       0.0280373      1.03367    0.5       0.135565       2.08139⌉
  0.95      0.482445       1.45839 '  0.95      4.11179        4.1132 ⌡
  0.99      2.69899        2.60364    0.99      34.18          4.19975
  Mean      20.9685        29.4023    Mean      127.689        15.3807
  Variance  4.1733×10^6    56.4701    Variance  5.16397×10^7   25.8577
```

Results from previous 2 analytic [U@.99] & 1 targeted [U@.95,.99] methods (denominators) for V@.99, compared to corresponding results from 2-D nested procedure:
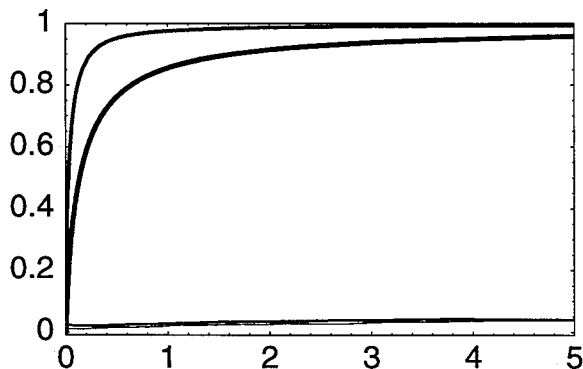
```
{34.2 / {4.1, .87}, {4.08 / 2.69, 32.9 / 34.2}}
```
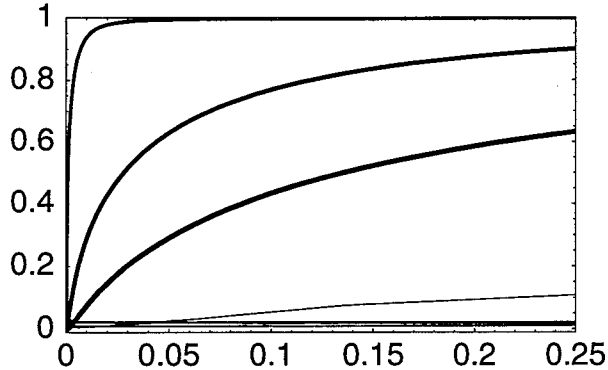
```
{{8.34146, 39.3103}, {1.51673, 0.961988}}
```

```
PlotCdf[Join[Rest[cvm], Rest[cdf]], Ymin → -.01, Xmin → -.0001, Xmax → 5];
```

```
PlotCdf[Join[cvm, cdf], Ymin → -.01, Xmin → -.0001, Xmax → .25];
```


end

end

Population Risk

expected cases for different values of total-population size, n, via the relation: <N> = n(<Rbar>). Note that Rbar (and hence <Rbar>) was scaled above by a factor of 10^6, and so needs to be rescaled by 10^{-6}.

```
npop = {100, 1000, 2000, 10^4, 30000, 10^5, 10^6, 0.997963 * 10^8};
{RbarAng, npop * (RbarAng * 10^-6)}
```

```
{0.915155, {0.0000915155, 0.000915155,
    0.00183031, 0.00915155, 0.0274546, 0.0915155, 0.915155, 91.3291}}
```

To obtain the likelihood P_0 of 0 people at risk associated with specified population sizes, first derive the probability mass function corresponding to Rbar, adjust to reflect the fact that its last element is artifically high due to how QUAnalyze defines this last element, and then use the adjusted pmf to calculate the complementary conditional Poisson likelihoods { P_0, $1-P_0$ } corresponding to specified population sizes (npop):

```
pmf = {10^-6 #[[1]], #[[2]]} & /@ Pmf[Rbar];
{ri, pi} = Transpose[pmf]; {Plus @@ pi, Take[pmf, -2]}
```

```
{1., {{0.00136677, 0.00049975}, {0.0101966, 0.00049975}}}
```

```
adjpmf = Transpose[(Drop[#, -1] & /@ {ri, pi}) {1, ──────────────}];
                                                    1 - Last[pi]
{ri, pi} = Transpose[adjpmf];
{Plus @@ pi, Take[adjpmf, -2]}
```

```
{1., {{0.000343409, 0.0005}, {0.00136677, 0.0005}}}
```

```
o = {npop, npop * 7.6 / 70, p0 = (Plus @@ (e^-#*ri *pi)) & /@ npop, 1 - p0, npop * RbarAng * 10^-6};
TBL[Prepend[Transpose[o], {n, nRes, Po, 1 - Po, "<N>"}]]
```

n	nRes	Po	1 - Po	<N>
100	10.8571	0.999913	0.0000867535	0.0000915155
1000	108.571	0.999423	0.000577328	0.000915155
2000	217.143	0.999166	0.000833686	0.00183031
10000	1085.71	0.998468	0.00153171	0.00915155
30000	3257.14	0.997568	0.00243198	0.0274546
100000	10857.1	0.995219	0.00478123	0.0915155
1000000	108571.	0.974127	0.0258733	0.915155
9.97965×10^7	1.0835×10^7	0.5	0.5	91.3292

To obtain the # people at risk associated with $P_0 = 0.5$, use the adjusted pmf to calculate the conditional Poisson likelihood corresponding to a likelihood of 0.50 (by manual numerical optimization [not shown]) :

```
Plus @@ (e^-0.997963*10^8 *ri *pi)
```

0.5

Thus, only if about **100 million people** were exposed would it be *more* **likely than not** that there would be 1 or more cases.

end

Appendix 3.I
Functions Used

■ *Mathematica*® functions

Note: The following *Mathematica* shorthand notation was used that is not included in the list of functions below:

a+b = a plus b

a-b = a minus b

a b = a*b = the product of a and b

a/b =a divided by b

a^b = a to the power of b

{a,b,c,...} = List[a,b,c,...] = a "list" (i.e., array, vector, or set) of elements a,b,c, ...

fxn/@{a,b,c,...} = Map[fxn, {a,b,c,...}] = a new list made by mapping (i.e., applying)
                the function fxn onto each member of the list {a,b,c,...}

? Cholesky

Cholesky[M] gives the Cholesky decomposition c of a symetric positive definite square matrix M (i.
   e., the lower triangular matrix c such that c c' = M), provided Det[M] does not equal zero.

? Dimensions

Dimensions[expr] gives a list of the dimensions of expr.
   Dimensions[expr, n] gives a list of the dimensions of expr down to level n.

? Drop

Drop[list, n] gives list with its first n elements dropped. Drop[list, -n] gives
   list with its last n elements dropped. Drop[list, {n}] gives list with its nth
   element dropped. Drop[list, {m, n}] gives list with elements m through n dropped.

? Flatten

Flatten[list] flattens out nested lists. Flatten[list, n]
   flattens to level n. Flatten[list, n, h] flattens subexpressions with head h.

? Last

Last[expr] gives the last element in expr.

? Length

Length[expr] gives the number of elements in expr.

? Log

Log[z] gives the natural logarithm of z (
   logarithm to base e). Log[b, z] gives the logarithm to base b.

? MapThread

MapThread[f, {{a1, a2, ... }, {b1, b2, ... }, ... }]
   gives {f[a1, b1, ...], f[a2, b2, ...], ... }. MapThread[f, {expr1,
   expr2, ... }, n] applies f to the parts of the expri at level n.

? Max

Max[x1, x2, ...] yields the numerically largest of the xi. Max[{x1,
   x2, ... }, {y1, ... }, ...] yields the largest element of any of the lists.

? Min

Min[x1, x2, ...] yields the numerically smallest of the xi. Min[{x1,
   x2, ... }, {y1, ... }, ...] yields the smallest element of any of the lists.

? Prepend

Prepend[expr, elem] gives expr with elem prepended.

? Range

Range[imax] generates the list {1, 2, ... , imax}. Range[imin, imax]
   generates the list {imin, ... , imax}. Range[imin, imax, di] uses step di.

? Solve

Solve[eqns, vars] attempts to solve an equation or set of equations for the variables vars. Solve[
   eqns, vars, elims] attempts to solve the equations for vars, eliminating the variables elims.

? Sort

Sort[list] sorts the elements of list into
   canonical order. Sort[list, p] sorts using the ordering function p.

? Take

Take[list, n] gives the first n elements of list. Take[list, -n] gives the
   last n elements of list. Take[list, {m, n}] gives elements m through n of list.

? Transpose

Transpose[list] transposes the first two levels in
   list. Transpose[list, {n1, n2, ... }] transposes list so that the
   levels 1, 2, ... in list correspond to levels n1, n2, ... in the result.

end

■ *RiskQ* 4.0 functions

<< RiskQ`;

? AverageCdf

AverageCdf[cdfs,options:] generates a cdf which is the exact average of the
   input list of cdfs and/or cmfs. By default, the input cdfs are equally weighted (
   i.e., all cdfs are assumed to be equally likely). Use Weights→weights to specify
   weights. Use TestCdf→False to suppress automatic CdfQ test of input cdfs. Use
   Approximate →n (or →xlist) to return an approximation of the true average cdf
   evaluated at n>1 equal abcissa intervals (or at the supplied list of abcissa values).

? Cdf

Cdf[x, options] returns a matrix representing a cdf (cumulative distribution funtion) from which
   x is (assumed to be) sampled if x is a vector. The first point is {xlo,0} where xlo is
   assumed to be Min[0,Min[x]] unless xlo<Min[x] is entered with Xmin→xlo. The last point
   is {xhi,1} where xhi is assumed to be Max[x] unless xhi≥Max[x] is entered with Xmax→
   xhi. Cdf estimates the cdf corresponding to n samples of a continous random variate,
   using linear interpolation. Use Weights→w to obtain a cdf based on weight-vector w
   corresponding to list (in which case the Xmax option is ignored). If x is a cmf or a
   pmf, a corresponding cdf is output. Use Pmf→True (or the alternative function Pmf) to
   obtain the probability mass function (pmf) corresponding to list. Use Simplify→False to
   suppress default distribution-simplification algorithm. To obtain the sample cdf (a step
   function) corresponding to list, or to model a discrete random variate, see Cmf. See also RQ.

? Data

Data[datarows, $expr_1$, ...] returns a list of data rows specified symbolically as a function of
   the input datarows list, where each $datarows_i$ = {x_{i1}, x_{i2},..., x_{in}} has n columns, and
   $expr_k$ are Data arguments. If datarows is a list but not a list of lists, then it is
   assumed to specify a single data column. By default, $datarows_1$ must be a list whose
   jth element ($name_j$) is a unique symbol or string used to name the variate whose values
   x_{ij} appear in the rest of the jth data column for j=1,...,n; however, if $expr_1$ is a
   vector containing n symbols and/or strings, then $expr_1$ j=$name_j$ is assumed. If $expr_1$
   is a non-Rule expression (e.g., involving any of the $name_j$), then $expr_1$ is returned
   evaluated using the specified data column(s). Otherwise, $expr_k$ must specify one or
   more of the following options (described below) to transform datarows: Append (or
   Replace), Classify (or Bin), Complement, Drop, Fill, Interpolate, Intersection, Merge,
   Names, Number, Rename, Restructure, Set, Shift, SortBy, Take, and/or Union. These
   options are applied in the order they appear (one or more times) in $expr_k$. Evaluate
   Data[option] to get information about any Data option. Column names (e.g., $name_1$)
   appearing in any of these options are assumed to be among those defined for (e.g.,
   as the 1st row of) datarows; any corresponding reassigned name (e.g., X after the
   assignment X=$name_1$ has been made) used should appear as an argument of HoldForm (e.g.,
   as HoldForm[X]). Data should be nested only if the nested expr is a rule or rule sequence.

? EV

EV[x, options] returns the arithmetic average of (e.g., a vector) x, or the expected value of
   x if x is a valid cdf, cmf or pmf. If x is a vector, Weights→w may be used to obtain
   the weighted average value corresponding to the weights-vector w applied to x. If x
   is a cdf with >2 evenly spaced ordinate values (i.e., evaluated at equal probability
   intervals) and Empirical→True, then the minimum and maximum abscissa values are ignored.

? Edf

Edf: See EvaluateCdf, RQ.

? EvaluateCdf

EvaluateCdf[cdf, x,complement:False] calculates the probability p that a
   random variate distributed as cdf is less than or equal to x, using linear
   interpolation. If a third argument, True, is included, the output probability is 1-
   p. The input x may be a list, in which case a corresponding output list is generated.

? FIT

FIT[xy, fxn_List, x_Symbol, options]   fits the General Linear Model (GLM), y(x) = Sum[qi Fi(x)]
  for i=1...np, to xj-yj data (for j=1...n) given in xy (an n-by-2 matrix) by direct or
  generalized least-squares regression, assuming yj are normally distributed as N(Eyj,
  Sqrt[v/wj]) with wj=1 by default and v estimated by the mean square of y-residuals (
  unless KnownVariances→ True is used, in which case yj ~ N(0,Sqrt[1/wj] is assumed).
  Use NYatX→nyj, with integer nyj>0 (ny=1 by default) or nyj an n-lengthed such list,
  to treat yj as means, in which case corresponding sample stand. devs. sj of nyj y-
  values must be specified using Errors→sj. Use Weights→wj to similarly specify known
  weights wj; or use Weights→{Wyhat, yhat, df} to specify that wj=(Wyhat| yhat=y(xj))
  or that wj=(Wyhatj|yhat=y(xj))--where Wyhat is an expression (or Wyhatj is a list of
  n expressions each) involving the symbol yhat--in which case the fit is obtained by
  iterative reweighting assuming df (=0 if not specified) extra degrees of freedom are lost
  in estimating Wyhat from the data (& use MaxIterations →maxit and/or Tolerance→tol to
  override defaults). If Report→True, SDs and 100p% conf. limits on ai, R^2, a chi-square
  test-of-fit, ANOVA table, F-tests of GLM-fit and nonzero qi for i>0, and a plot are
  all printed (use Report→Plot to add a plot). The qi estimates are output, along with:
  covariance matrix, the list {xval,yhat,yLCL, yUCL}, a sum-of-squares & assoc. degree-of-
  freedom matrix, the F-values and their p-values, the chi-square value and its p-value,
  the fitted function, and/or a plot using Output→{CV, xval, SumSquares, F, PvalF, X2,
  PvalX2, BestFit, [and/or] Plot}, where xval may be a list. Use Confidence→p to change
  p from 0.95, and use Xmin→xlo, etc.(see PlotData Options) to change plot defaults.

? Idf

Idf: See InverseCdf, RQ.

? InverseCdf

InverseCdf[cdf, p, options] evaluates cdf at the cumulative probability value p, for any
  valid cdf or cmf. The input p may be a list of probability values, in which case a
  corresponding output list is generated. Use TestCdf->False to suppress default CdfQ test.

? NormalCdf

NormalCdf[z, s:, n:100] = the standard Gaussian cdf; i.e. the probability p that Z<=z for
  real z and standard normal random variate Z. If z is a list, p is the corresponding
  list. If s is set to Inv, then the inverse standard Gaussian cdf is returned for
  argument(s) z where 0<=z<=1. If s is entered as a nonegative real number, then an
  approximate cdf is returned corresponding to the parameters {z= mean, s= stand. dev.}
  for a nonstandard random variate Z, evaluated at n equiprobable quantile intervals.
  In evaluating the inv. stand. Gaussian cdf for Min[p,1-p]>2.21 10^-7, NormalCdf makes
  use of an 11th-order polynomial approximation with an absolute error < 0.503 10^-6.

? PlotCdf

PlotCdf[cdfs] returns a plot of a cdf or
  of several cdfs entered as a list of cdfs. See PlotOptions.

? PlotData

PlotData[data,options:]   plots an N-by-2 (or (x vs y) data set (DS), or a list of n such sets,
  with points joined by lines (unless JoinPoints->False is used). Change point style with
  Style->list which by defualt is {OO,OA,OB,OV,OD,O,A,B,V,D} = {open Point, Triangle,
  Square, InvTriangle, Diamond,...(& their solid equivs.)}; use {TO,TA,TB,TV,TD} for
  transparent open symbols; use {P,X,M,I} for {plus,cross,dash,bar}; & use J to join
  points from adjacent DSs. Size and JoinPoints may be n-lengthed lists, where n<=Length[
  Style] <=2n-1 depending on how many Js are in Style; JoinPoints->False is enforced for
  DSs referenced by J in Style.Use FitTo->{f[x],x} to include a plot of f[x] (which may
  be a list of functions) vs. x. Use data=Plot to plot functions only. See PlotOptions.

? QUAnalyze

QUAnalyze[cdfs, Fxn, nsam, nsim, options:] performs a quantitative uncertainty analysis involving
simulated values Fxn_{jk} of $Fxn[var_1,...,var_n]$, where Fxn is a user-defined listable
function, $j = 1,...,$ nsam, and $k = 1, ...,$ nsim. Uncertainties in var_i, $i=1,...,n$,
are specified by the corresponding input cumulative probability distributions, cdfs =
$\{cdf_1, ..., cdf_n\}$, where each cdf_i must be either a valid Cdf object (for which TrueQ[
cdf_i]==True) or a valid symbolic cdf-specification (see SimulateCdf). All cdfs are
by default uncorrelated, unless Correlate→T is used to specify T as the target rank-
correlation matrix (or as its upper-right rows--see Reflect). The list {SimReport,
cdfFxn, cvmFxn} is output, where: SimReport lists the coefficients of variation (as a %)
of $EV[Fxn_j]_k$ and corresponding p-fractiles of Fxn, the maximum of Jennrich chi-square
values assessing homogeneity with T, its degrees of freedom, and the corrsponding Hommel-
adjusted p-value; cdfFxn characterizes Fxn uncertainty (as the means of nsim sorted
sets of nsam sample values of Fxn--i.e., as nsam mean Fxn-fractile values--where nsam >
n and nsam>nsim>1); and cvmFxn lists the corresponding coefficients of variation of
the nsam fractile means (and so summarizes corresponding Monte-Carlo sampling error).
By default, the minimum and maximum possible values of Fxn are assumed to be xlo = Min[
Fxn_{jk}] and xhi = Max[Fxn_{jk}], respectively; use Xmin→xmin and/or Xmax→xmax to change
these defaults (provided xmin<xlo and xmax>xhi). Use Fractiles → p to specify the list
of p-fractiles of Fxn to be used to summarize simulation quality. Use SimIn→ML to
specify the cdf-simulation values to be used, where either: (1) ML is a list of nsim
matrices each n-by-nsam in dimension (as output by SimulateCdf); or (2) ML is a list
of nsim elements each of the form {M, RankCorrelations→R, Jennrich→{x2,df,p}} (i.e.,
each element of ML is a list of the form output by SimulateCdf using the Report→Append
option), where M is an n-by-nsam matrix, R is an n-by-n matrix, and x2, df, and p are
numbers with 0<p<1). If SimIn is specified, the QUAnalyze cdfs argument supplied may be
the integer n. Note that cdf and cvm may be plotted together because they use common
abscissa values, which are scaled by an n-fold factor if the option Scale→n is used.

? Reflect

Reflect[upper, diagvec:Automatic, anti:False] returns an n-by-n symetric square matrix M given
upper, an (n-1)-length list of (n-1-i)-length lists (i=1,...,n) that represent the first n-1
rows of upper elements (without the diagonal elements) of M. The diagonal is a vector of ones
by default, or may be entered as the 2nd argument (either a constant or an n-length list).
If the 3rd argument is set to True, then the corresponding antisymetric matrix is returned.

? RQ

RQ[operation, distribution, parameter(s), z] performs an operation Cdf (=cumulative
distribution function), E=M= Mean (=expected value), V=Var=Variance, D=Range=Domain,
P=Pr=Prob=PDF, C=Edf=CDF (CC= C complement), Q=Idf=Quantile (QC= Q complement),
or Test= (test validity of 2nd & 3rd RQ parameters) on a B=Beta, Bi=Binomial, X2=
ChiSquare, E=Exponential, F=FRatio, G=Gamma, Geo=Geometric, H=Hypergeometric, LN=
LogNormal, Lg=Logistic, NBi=NegativeBinomial, N=Normal, Psn=Poisson (=P), T=StudentT,
Tri=Triangular, U=Uniform, W=Weibull (=Wbl), or M=Empirical (=Cdf=Cmf) distribution
with the specified parameter value(s) or for the particular cdf/cmf, at the point(s)
z. If z is included with the Cdf operation, the output cdf is given for z+1 points.

? SimplifyCdf

SimplifyCdf[cdf] returns any valid input cdf or cmf in its simplest
possible form, that is, without any unnecessarily repeated or redundant elements.

? SimulateCdf

SimulateCdf[cdf(s), nsim, options:]   generates a list of nsim values simulated from an input
   cdf, or of n lists of nsim values with the ith list simulated from the ith of an
   input list of n cdf objects with a target rank- correlation matrix T. Input T using
   Correlate→T for a square matrix T (or its upper-right rows--see Reflect); by default
   T is an identity matrix. Each cdf must be either empirical (such that TrueQ[cdf]==
   True) or a valid {type, par} cdf-specification (see RQ). Simulations use a Systematic
   Latin Hypercube (SLH) method, adjusted (unless Correlate→ False is used) to yield
   variates whose true rank-correlation matrix R approximates T. Alternative methods
   may be specified with SimMethod →LatinHypercube (→LH) or →Random (→U,→Uniform). If
   the first argument is entered as n for n>0, then cdf(s) are assumed to be n standard
   normal cdfs and T=R is the actual product-moment correlation matrix. Note that unless
   Correlate→False, nsim must be >n. Use Report→False to suppress the Jennrich- function
   report comparing T vs.R (suppressed by default for normal variates, for which T=R), or
   use Report→Append to append R and {chi2, df, pval} from this report to the output (see
   Jennrich). Use SimIn→ mymatrix (SimOut→True) to use an input (output the simulated)
   rank-matrix. Use TestCdf→False to save time if CdfQ[cdf]==True for each input cdf.

? StandardizeCdf

StandardizeCdf[incdf, Values, options:] returns a new cdf based on linear interpolation
   of incdf (any cdf or cmf) evaluated at Values, where Values are assumed to be
   probability values, except that Values are treated as cdf abcissa values if any of
   the Values are <0 or >1 or if the option ProbabilityValues→ False is used. Values
   must be either a list or an integer>0; in the latter case cdf-evaluation occurs in n
   equal increments over the specified range of probability or abcissa values. If the
   Midpoints→True option is used,then cdf-evaluation occurs at the midpoints of the
   successive element-pairs in the specified set of values, rather than at those values
   themselves. If incdf is a list of cdfs, then a corresponding list of standardized
   cdfs is output. Use TestCdf→False to suppress automatic CdfQ test of input cdfs.

? TBL

TBL[x] = TableForm[x, TableSpacing->1]. TBL[x,n] = TableForm[
   x, TableSpacing->n]. TBL[x,n,r] = TableForm[Take[x,r], TableSpacing->n].

? WriteMatrix

WriteMatrix[filenameString, dataMatrix, separatorString] writes a data matrix to
   the specified filename as an ASCII file. The separatorString is a tab by default.

end

■ Other Functions

? LSMin

LSMin[x,y,p,fxp,options] Attempts to reduce x2, i.e., the chi-square (i.e., weighted sum of squared residuals) between a list y of data and a user-defined function fxp[x,P] of corresponding independent-values x and parameters P, starting with the initially guessed parameter list p, returning {phat,sd,{x2,df,p}} where phat is the list of asymptotic maximum-likelihood parameter-value estimates, sd is the corresponding list of standard deviations (or the full covariance matrix if Output->CVM is used), x2 is the goodness-of-fit chi-square value, df=(Length[y] - # est. parameters), and p is the corresponding p-value. Use NYatX->ny if the y-values are the means of (a list of) ny corresponding values, with corresponding standard deviations sdy all 1 (unless SDY->sy is used). It is assumed that yj~N(Eyj,Sqrt[v/wj]) with wj=1 by default (and v = mean square of y-residuals if KnownVariances-> False is used, in which case p is meaningless; otherwise v=1). Use Weights->wj to specify weights wj. Use Weights->{Wyhat, yhat, df} to specify wj=(Wyhat| yhat=y(xj)) for Wyhat an expression (or list of n expressions each) involving the symbol yhat, in which cases the fit is obtained by iterative reweighting. If weights are not specified, KnownVariances-> False is assumed. Use Parameters-> pinlist with ordered integer index-list pinlist to restrict optimization to a subset of p specified by pinlist. The search stops if reductions in chi-square become less than tol=0.001 (reset using Tolerance->tol) or if iterations > maxit=100 (reset using Maxit->maxit; output appended with 'Warning'). Set Progress->True to see intermediate output (at precision p using SeePrecision-> p). Levenberg-Marquardt minimization of the chi-square objective function is used (WH Press et al., Numerical Recipies, Cambridge U. Press, New York, 1986, pp. 521-528), with shifts at each step having a relative size equal to 1000 (reset using Step->size). In the case of unknown sigy, generalized (i.e., iteratively reweighted) x2-minimization is performed (see Carrol and Rupert, Transformation and Weighting in Regression, Chapman and Hall, New York, 1988). Needs MarqCof, Partial, Bracket, ParaMin, and Mathematica's CDF and ChiSquareDistribution functions.

? MSDx

MSDx[GMx,GSDx] returns the arithmetic mean and arithmetic sdandard deviation of a lognormal variate X that also has the specified geometric mean GMx and geometric sdandard deviation GSDx, based on the method of moments.

? GMGSDx

GMGSDx[Mx,SDx] returns the geometric mean and geometric sdandard deviation of a lognormal variate X that also has the specified arithmetic mean Mx and arithmetic sdandard deviation SDx, based on the method of moments.

? GMGSDx1

GMGSDx1[cvWant,cv2] returns the GM and GSD of a lognormal variate X1, such that the product X1*X2 has the desired coefficient of variation (CV) = cvWant, conditional on the lognormal variate X2 having an arithmetic mean and CV equal to 1 and cv2, respectively, based on the method of moments.